SEISMICA

JUNE 2024 | VOL 3 | NO 1

ISSN 2816-9387

McGill Library Bibliothèque

Issue 3, Volume 1, 2024

Aims and Scope

Seismica publishes original, novel, peer-reviewed research in the fields of seismology, earthquake science, and related disciplines. Seismica is a community-driven, diamond open-access journal. Articles are free to publish and free to read without a subscription, and authors retain full copyright.

Seismica strives to be **accessible**, **transparent**, **respectful**, **credible**, and **progressive**.

Seismica's scope includes a wide range of topics in seismological and earthquake sciences. A non-exhaustive list of topics that fall within the scope of Seismica includes: Fault-slip and earthquake source phenomena, Earthquake records, Imaging the Earth, Theoretical and computational seismology, Beyond Earth-tectonic applications, Techniques and instrumentation, Earthquake engineering and engineering seismology, Community engagement, communication and outreach.

For submission instructions, subscription and all other information visit: https://seismica.library.mcgill.ca/.
For editorial questions or issues, please contact info@seismica.org.
For technical assistance, please contact info@seismica.org.
For appeals, please contact appeals@seismica.org.

About the cover

Cover Caption Drone-based lidar platform preparing to land after conducting a survey near Shǟr Ndü Chù (Duke River), Yukon, Canada, as part of a broader project coordinated by the Yukon Geological Survey to study the neotectonics and geothermal potential of the Eastern Denali fault. The drone offers a cost-effective way of obtaining otherwise expensive airborne lidar data, and compares favorably against established methods of topographic mapping, allowing landscapes to be surveyed in finer detail than was previously possible. This technology is particularly important in forested regions where dense vegetation would otherwise obscure subtle landforms, such as those produced by crustal faults in low strain settings. Credit: Guy Salomon.

Editorial Board

We warmly thank the following volunteers of the Editorial Board that have been involved in the issue.

Volunteers

We warmly thank the following volunteers that have been involved in the issue.

Issue 3, Volume 1, 2024

Table of Contents

Articles

Reports (excl. Fast Reports)

doi:10.26443/seismica.v3i1.1190

SEISMICA

Extensional failure in a weak slab under slab pull – the 2023 M_w 6.4 Quiché, Guatemala, earthquake

Timothy J. Craig * [1], **Amber Hull** [2]

[1]COMET, Institute for Geophysics and Tectonics, School of Earth and Environment, University of Leeds, Leeds, UK, [2]Institute for Geophysics and Tectonics, School of Earth and Environment, University of Leeds, Leeds, UK

Author contributions: *Conceptualization*: TJC. *Formal Analysis*: TJC, AH. *Investigation*: TJC, AH. *Writing - original draft*: TJC. *Writing - Review & Editing*: TJC, AH. *Supervision*: TJC. *Funding acquisition*: TJC.

Abstract The 2023 M_w 6.4 Quiché earthquake is the deepest recorded major ($M_w > 6$) earthquake to have occurred in the Cocos slab beneath Central America, at a depth of $\sim$ 255 km. Here, we refine the source parameters of both the Quiché earthquake, and the only other event at comparable depths (the 1997 M_w 5.5 Jutiapa earthquake), confirming both their exceptional depth within the downgoing slab, and their down-dip extensional mechanism. That the Cocos slab remains capable of hosting major intraslab earthquakes, with mechanisms consistent with down-dip extension, near, or at, the tip of the contiguous slab, suggests that the slab itself is weak, such that the minimal stresses derived from supporting the negative buoyancy of the short section of slab down-dip from this earthquake are still sufficient to lead to brittle failure of the slab.

Resumen El terremoto de Quiché de magnitud M_w 6.4 en 2023 es el terremoto de mayor magnitud ($M_w > 6$) registrado en la zona más profunda de la placa de Cocos bajo América Central, a una profundidad de $\sim$ 255 km. Aquí, refinamos los parámetros fuente tanto para el terremoto de Quiché como para el único otro evento a profundidades comparables (el terremoto de Jutiapa de magnitud M_w 5.5 en 1997), confirmando tanto su profundidad excepcional dentro de la placa descendente, como su consistente mecanismo extensional hacia abajo. Que la placa de Cocos siga siendo capaz de experimentar terremotos intra-slab de gran magnitud, con consistentes mecanismos de extensión hacia abajo, cerca o en el borde con la placa contigua, sugiere que la placa misma es débil, a tal punto que las tensiones mínimas derivadas de este terremoto, asociadas a la flotabilidad negativa de la pequeña sección de la placa descendente, siguen siendo suficientes para producir el fracturamiento frágil de la placa.

Production Editor:
Gareth Funning
Handling Editor:
Marino Protti
Copy & Layout Editor:
Théa Ragon

Received:
January 24, 2024
Accepted:
March 22, 2024
Published:
May 28, 2024

1 Introduction

The oceanic Cocos plate subducts beneath Central America along the Middle America Trench, giving rise to both widespread seismicity on the subduction megathrust and to prolific (although unevenly distributed) seismicity within the Cocos slab as it descends into the upper mantle. Although contributing only a small proportion of the overall moment release associated with the Central American subduction zone, intraslab events can, on occasion, be both large and damaging. In the case of Central America, these intraslab events include the damaging M_w7.7 2001 El Salvador earthquake (Vallée et al., 2003), the M_w7.4 1999 Oaxaca/Tehuacán earthquake (Singh et al., 2000), and the 2017 M_w8.2 Tehuantepec and M_w7.1 Puebla earthquakes (Melgar et al., 2018a,b).

In its northern sections, the Central American slab is dominated by the flat slab region under southern Mexico (e.g., Kim et al., 2010; Manea et al., 2017). East of $\sim$ 97° W, the slab transitions via a region of probable slab tearing (e.g., Rogers et al., 2002; Manea et al., 2013) to a more classical slab geometry, dipping gently down into the upper mantle (e.g., Syracuse et al., 2008; Manea et al., 2013). East of the flat slab region, the slab shows a fairly consistent geometry, characterised by its dip gradually increasing from $< 20°$ to $> 60°$ at fairly consistent slab curvatures (Hayes et al., 2018). Current slab models suggest the contiguous downdip slab extends to depths of $\sim$ 250-300 km — below this depth, the nature of the slab becomes unclear, with different data suggesting either a gap between the shallow slab and a detached slab in the mid mantle (Rogers et al., 2002) or a fragmentary, perforated, slab subject to through-going mantle flow (Zhu et al., 2020; Xue et al., 2023).

The causative rheological mechanism allowing brittle failure in such intraslab earthquakes to occur remains uncertain, with both dehydration embrittlement and shear-heating driven thermal instability remaining as viable candidates (e.g., Hosseinzadehsabeti et al., 2021; Wimpenny et al., 2023; Prakash et al., 2023). However, no matter what the rheological control allowing seismogenic failure is, the consistency and regional coherency of deformation illuminated by intraslab seismicity requires that their spatial occurrence is controlled by the intraslab stress state – well established to be a function of the interplay between stresses related

*Corresponding author: t.j.craig@leeds.ac.uk

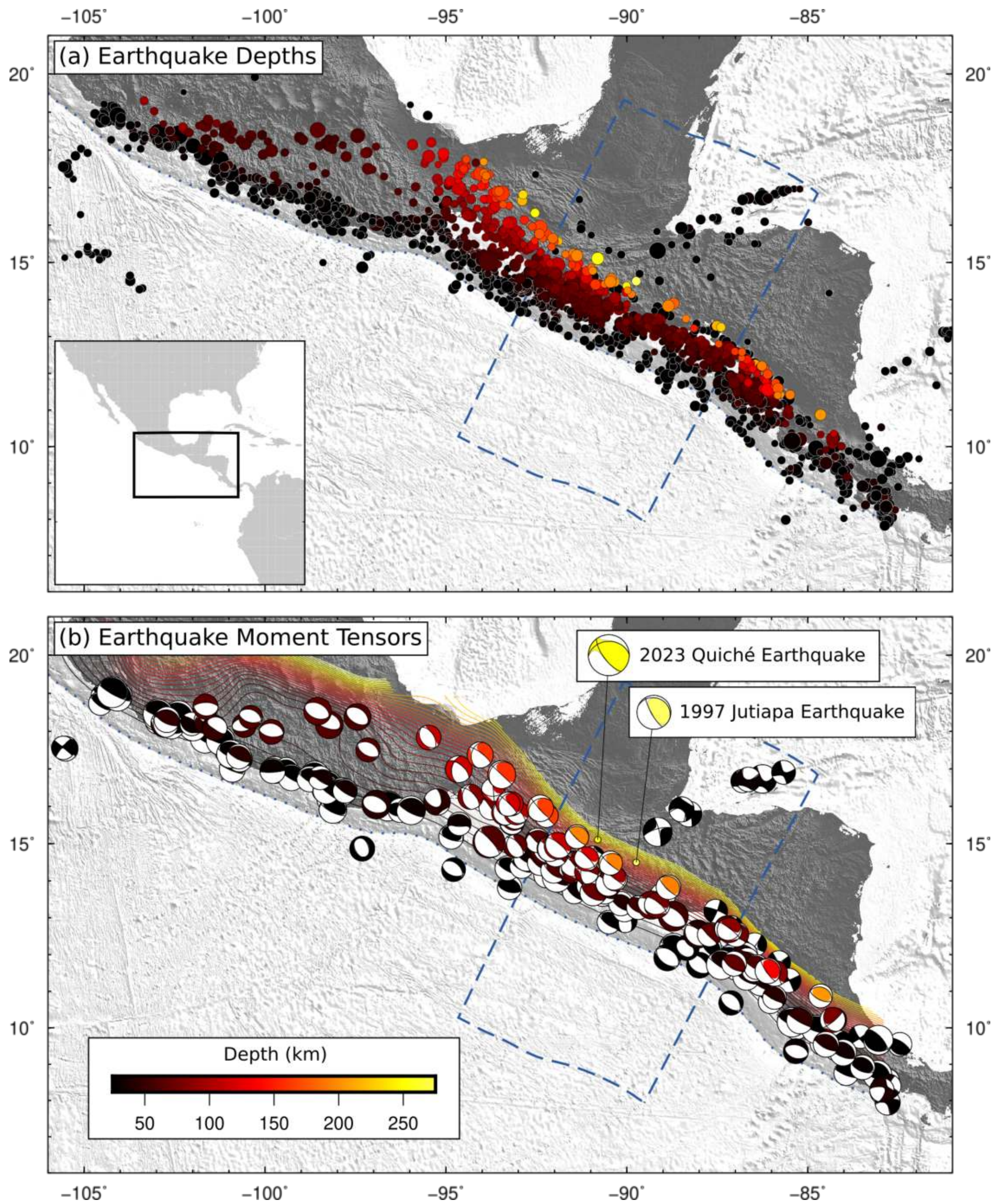


Figure 1 Earthquakes along the Central American Subduction Zone. (a) Earthquake depths. Inset panel shows a regional location map; (b) Earthquake moment tensors, from the gCMT catalogue. Compressional quadrants are shaded by earthquake depth. Inset boxes highlight the 2023 Quiché and 1997 Jutiapa earthquakes, and use our moment tensors and locations. Contours show the Slab2 model (Hayes et al., 2018). Dashed blue box shows the region used in plotting the cross sections shown in Figure 5.

to the negative buoyancy of the slab (slab pull), tractions on the edge of the slab from its interaction with the surrounding mantle, stresses relating to the bending and unbending of the slab, and stresses arising from the evolving thermochemical state of the slab (e.g., thermal expansion, volumetric changes resulting from mineralogical phase transitions) (e.g., Manea et al., 2006; Bailey et al., 2012). The dominant two are believed to be slab pull, which plays a role in driving global plate tectonics, and the bending stresses. Which of these dominates the overall stress state of the slab likely varies between different geodynamic settings (Sandiford et al., 2020; Craig et al., 2022; Sippl et al., 2022).

This study focuses on the M_w 6.4 2023 Quiché earthquake (Figure 1), which occurred beneath the central Guatemalan cordillera on the 17^{th} May 2023, at 23:02:00 (UTC). Preliminary locations placed this earthquake at depths of ~ 255 km (as reported by the NEIC; please see Table 1), at the leading edge of the contiguous section of the subducting Cocos slab beneath Central America. The depth and location of this earthquake make it stand out against the backdrop of other seismicity associated with the Central American slab – both as a comparatively large-magnitude event for the Central American slab, but also as one at substantially greater depths than generally recorded for earthquakes in the slab beneath

Origin time (UTC)		Method	Lat (°)	Long (°)	Depth (km)	M_w	M_{rr}	M_{tt}	M_{pp}	M_{rt}	M_{rp}	M_{tp}	γ
2023/05/17	23:02:00.5	NEIC	15.181	-90.815	253.5	6.4	-	-	-	-	-	-	-
	23:02:03.2	gCMT	15.17	-90.99	253.0	6.4	0.305	-0.334	0.030	0.326	-0.525	0.124	84%
	23:02:05.6	This Study	14.82‡	-91.12‡	256.0	6.4	0.449	-0.375	-0.074	0.222	-0.428	0.306	94%
1997/05/15	04:39:21.5	NEIC	14.460	-89.775	274.2	4.9†	-	-	-	-	-	-	-
	04:39:23.3	ISC-EHB	14.490	-89.741	279.4	4.9†	-	-	-	-	-	-	-
	04:39:26.3	gCMT	14.53	-89.85	272.9	5.5	0.170	-0.072	-0.073	0.315	-0.616	0.043	92%
	04:39:25.6	This Study	14.12‡	-89.82‡	270.0	5.5	0.328	-0.213	-0.115	0.276	-0.568	0.136	85%

Table 1 Earthquake source parameters from different seismic catalogues. Note that at the time of writing no solution is yet available for the 2023 Quiché earthquake from the ISC. Magnitudes denoted † are instead m_b rather than M_w. Locations marked with ‡ are unreliable.

Central America. Along with the 2023 Quiché earthquake, we also revisit a M_w 5.5 earthquake from 15th May 1997, which occurred under Jutiapa in southeastern Guatemala, $\sim$150 km ESE along the strike of the subduction zone from the Quiché earthquake, but at a similar depth – also at the leading edge of the contiguous slab. Together, these two earthquakes represent the two deepest larger-magnitude events, and define the tip of the seismogenic slab.

Here, we present a refinement of the seismologically-determined source parameters of these two earthquakes, consider their geodynamic context within the Cocos plate, and the implications of such deep intraslab earthquakes for the force-balance of the subducting plate.

2 Earthquake source parameter determination

We refine initial estimates of the source parameters of the Quiché and Jutiapa earthquakes using global seismic data. We use the Grond software of Heimann et al. (2018) to invert seismological waveform data for the earthquake location (latitude, longitude, depth), source duration, magnitude, and six components of the moment tensor. We draw on both vertical and horizontal component data recorded at teleseismic distances ($30°$ – $90°$) from the earthquake epicentre. Horizontal components are rotated into earthquake-relative radial and transverse components. Station response functions are removed from all seismograms, and data are filtered to frequencies of 0.025 – 0.25 Hz using a four-pole Butterworth bandpass.

We use vertical component data to invert for the direct *P*-wave and its associated depth phases, and both radial and transverse data to invert for the *S*-wave and its associated depth phases. Misfits from radial and transverse component waveforms are downweighted during inversion by a factor of two, to reduce over-fitting of the usually higher-amplitude *S*-wave and its depth phases. Inversion windows are taken from 20 seconds before the predicted onset of each direct phase, to 120 after – a time range estimated to encompass the direct arrival and principal depths phases (*pP, sP, pS, sS*), based on the initial catalogue depth, and verified visually. Whilst the majority of similar studies use only vertical and transverse component data, we find here that the additional use of radial component data, in this case, makes a minor improvement in the resolution of the source mechanism, due to the inclusion of the *pS* depth phase, particularly for the Quiché event.

We invert for the six components of the moment tensor (constrained to be purely deviatoric with no isotropic component), along with three location parameters, moment, and source duration. Synthetic and observed waveforms are realigned during each iteration using on a time shift which maximises the cross correlation value between the observed and synthetic traces at each station and for each component independently.

Information about the velocity structure of the overriding Central American plate is limited, in comparison to other subduction zones, due to the relative sparsity of near-field instrument deployments. As a result, we use a velocity structure based on the global ak135 velocity model (Kennett et al., 1995). Given the inclusion of a waveform-realignment step in the inversion approach, our modelling is most sensitive to the velocity structure at depths between the earthquake source and the free surface. This being the case, our results are not particularly impacted by the lack of a fast, cold mid-mantle slab in ak135, and this does not have a significant impact on the determination of either source depth or source mechanism. However, teleseismic source inversions, especially in cases where station-specific temporal realignment is included, are typically insensitive to small changes in the lateral location of the earthquake. As such, although we do allow our inversion to re-determine source latitude and longitude (see Table 1), we note that these are poorly constrained, with substantial variability in the range of acceptable solutions, and therefore we consider these parameters to be unreliable.

Figures 2 and 3 show inversion results for the Quiché and Jutiapa earthquakes respectively, showing example radial, transverse and vertical component waveforms, the resultant probabilistic moment tensor, and parameter histograms for depth, strike, dip, rake, and the degree to which the mechanism contains any non-double couple (nDC) component. In this case, we assume that a well-constrained mechanism would be a pure double couple, and that any nDC component would reflect the mapping of noise into the solution. For both events, the nDC moment required is only a small fraction of the overall moment release required, indicating that

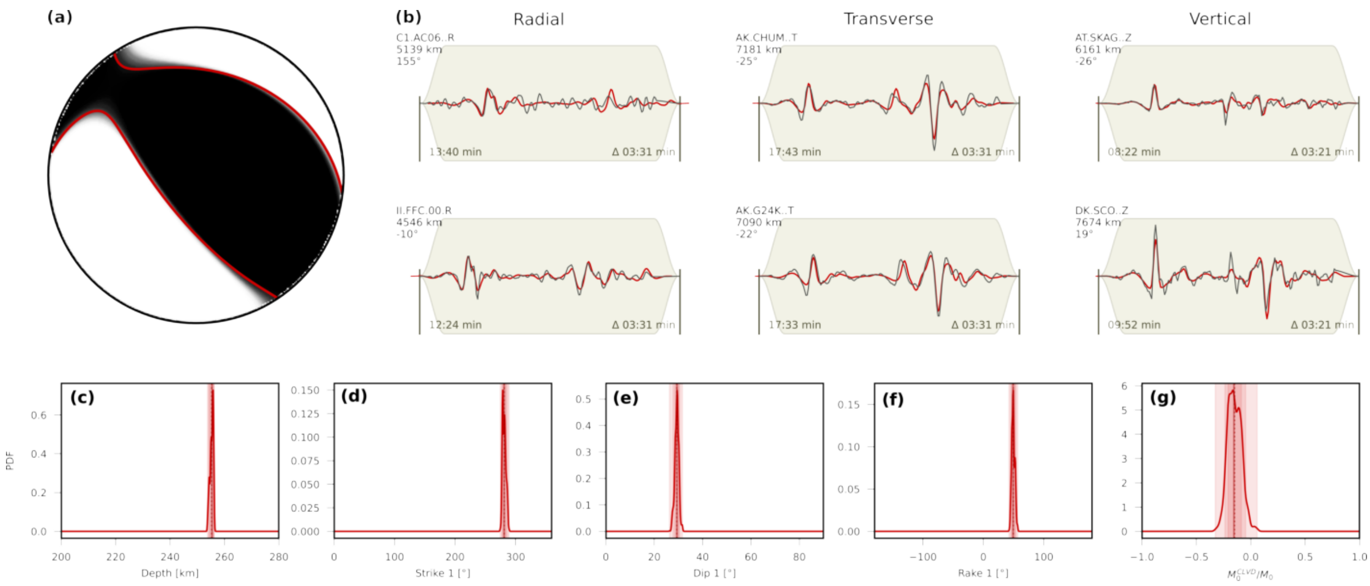


Figure 2 Earthquake source determination results for the 2023 Quiché earthquake. (a) Probabilistic moment tensor. Red lines show the minimum misfit moment tensor. (b) Observed waveforms (black) and calculated synthetics (red) for the minimum misfit solution. Shown are 6 examples traces, two each for radial, transverse and vertical components. Note that the instruments shown for each component varies. Annotations give the network and station name, epicentral distance, azimuth, trace start time, and inversion window length. (c) – (g) show probability density functions for depth, strike, dip, rake, and the ratio of the moment proportions allocated as a compensated linear vector dipole (CLVD) to the overall moment.

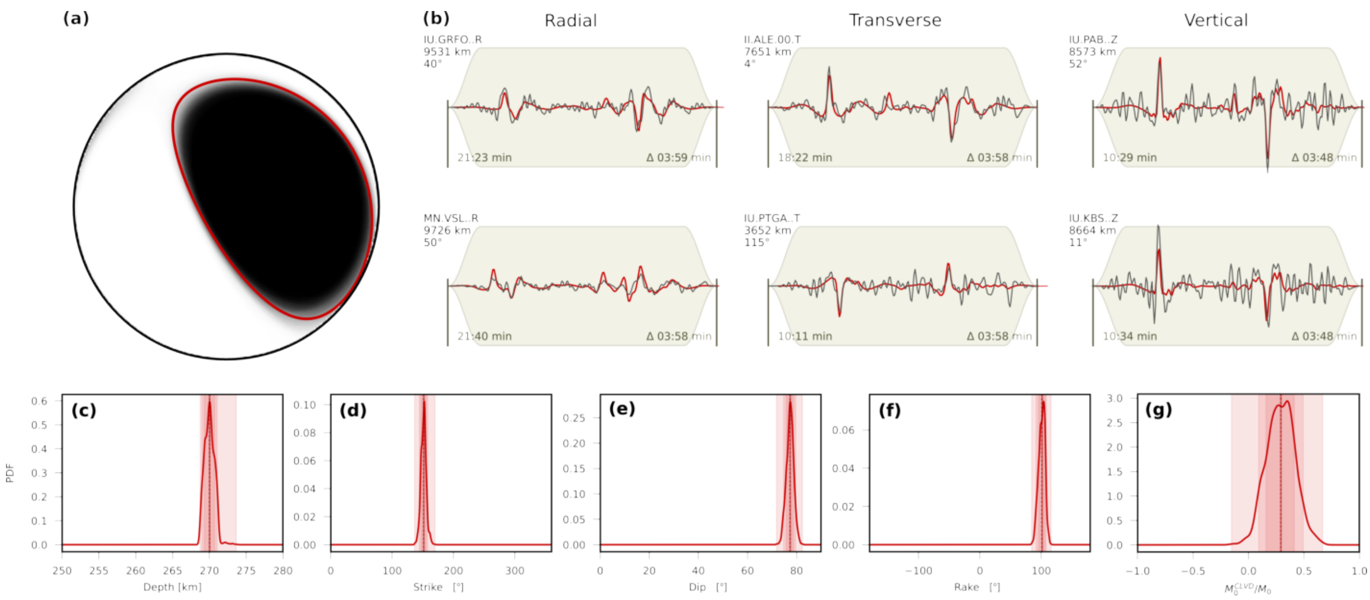


Figure 3 Earthquake source determination results for the 19997 Jutiapa earthquake. Panels are as shown in Figure 2.

our solutions are reasonably well constrained, and free from significant noise influence – a conclusion visually supported by inspection of the waveforms in Figures 2 and 3, which shows clear, relatively noise-free phase arrivals.

As the waveform fits in Figures 2 and 3 show, we are able to fit both the timing and amplitude of multiple phases across the waveform sections used. Both of our two earthquakes yield well-constrained depths, with Jutiapa being slightly deeper (270 km) than the Quiché event (256 km). These events are verified to be the deepest significant earthquakes within the Cocos plate yet recorded using global seismic data.

Whilst the two earthquakes differ slightly in their mechanism, due to the different signs of their small nDC components, the double couple component of their respective mechanisms is overall similar, representing almost pure dip-slip faulting, striking marginally obliquely to the strike direction of the slab, with a steeply-dipping southeast striking nodal plane, and a shallowly-dipping northwest striking nodal plane. Mechanism orientation parameters in all cases are extremely well constrained, and are consistent with the orientation of faulting within the Cocos slab in other earthquakes deeper than ~ 150 km.

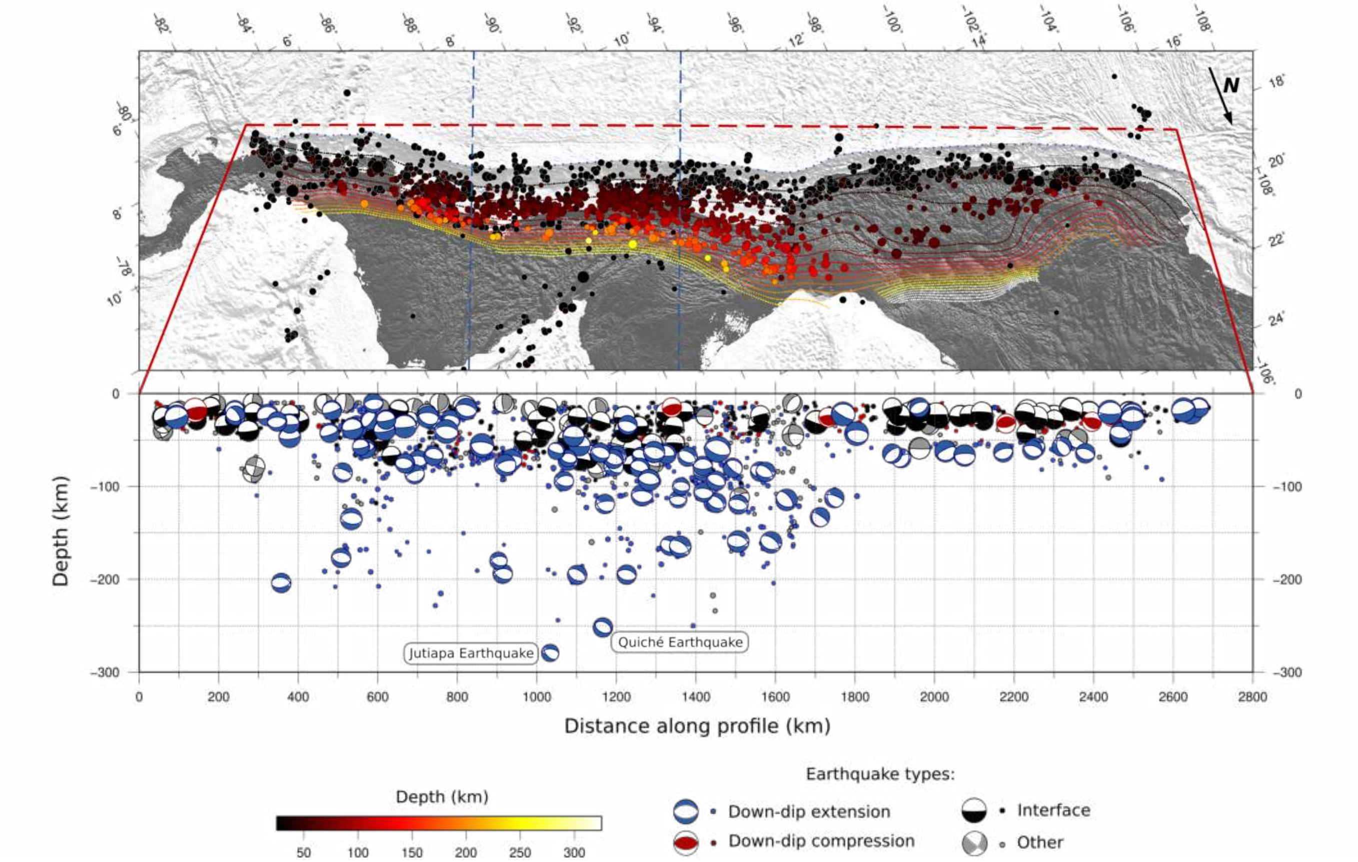


Figure 4 Reprojected earthquake moment tensors in the Central American Subduction Zone. Upper panel shows earthquake locations, coloured by depth. Contours show the slab model of (Hayes et al., 2018). Blue lines show the limits of the cross section shown in Figure 5, red line shows the projection line for the lower panel. Lower panel shows earthquake locations as a function of depth and distance along the projection line shown in the upper panel. Moment tensors (for earthquakes with $M_w \geq 6.0$ are rotated in three dimensions to be in a slab-relative reference frame appropriate for the location of the earthquake and its local slab geometry. The moment tensor for the 1997 Jutiapa earthquake is plotted despite having an M_w of 5.5.

3 Seismicity with the Central American Slab

In Figure 4, we show subduction-related seismicity along the Central American subduction zone, reprojected into a slab-relative reference frame such that the focal mechanisms shown are relative to the local slab surface from the Slab2 model of Hayes et al. (2018). Figure 5 shows a cross section through the slab beneath Guatemala, including a reprojection into slab-relative coordinates in Figure 5b.

Although in map view (e.g. Figure 1), many of the deeper earthquakes within the Cocos slab appear to be thrust-faulting earthquakes, when considered in a slab-relative reference frame, earthquakes within the slab instead almost entirely reflect down-dip extension within the slab. As Figure 5 demonstrates, intraplate faulting is dominated by downdip extension almost perfectly aligned to the direction of the slab dip. The slight mis-alignment between the strike of active intraslab faults and the slab strike (visible on Figure 4 and Figure 5a) probably reflects the slight misalignment between the trench orientation and the relict fabric of the incoming Cocos plate, reactivated in the outer rise (e.g., Masson, 1991; Ranero et al., 2005), and which likely remains active in the intraslab environment (Boneh et al., 2019).

The majority of slabs globally show a diversity of intraslab focal mechanisms, with a mix of both downdip compression and downdip extension, typically separated into discrete planes (double seismic zones) within the slab (Isacks and Molnar, 1969; Sandiford et al., 2019, 2020; Craig et al., 2022; Sippl et al., 2022). However, as we see from Figures 4 and 5, the Cocos slab stands out against this trend – almost all of the intraslab seismicity below ~75 km depth reflects down-dip extensional stresses. As Figure 5a shows these earthquakes show a remarkable degree of consistency in their slab relative mechanism orientations – a result of a relatively simple intraslab stress field, presumably dominated by slab pull, and with minimal impact from bending-related stresses after the initial phase of post-subduction unbending, reflected in the simple slab geometry seen beneath Guatemala on Figure 5c.

However, of particular note here is that both the Quiché and Jutiapa earthquakes occur at the very leading edge of the contiguous Cocos slab, in a region where the slab is subject to neither substantial bending-related stresses, nor to major buoyancy forces relating to the presence of a substantial high-density down-dip slab. Despite this, both earthquakes show complete consis-

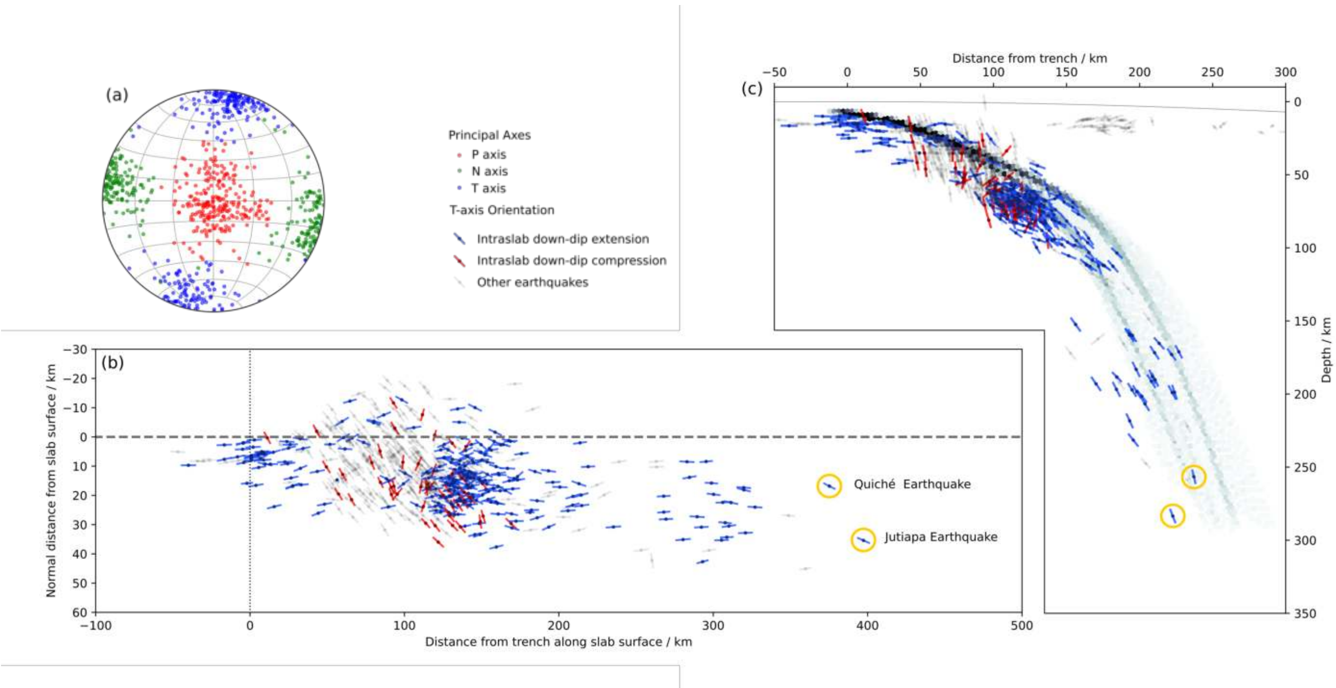


Figure 5 (a) Slab-relative orientations of the principal axes for the population of earthquakes consistent with down-dip extension. (b) Trench-perpendicular cross section for the region between the blue lines on Figure 1. Earthquakes locations are given by circular points, underlain by bars indicating the inclination of the T axis. Bars are coloured blue for those events consistent with down-dip extension, red for those consistent with down-dip compression, and grey for all other earthquakes. Locations are reprojected to show trench-perpendicular distance relative to the local trench. Background shows the variation in slab geometries across the region. Yellow circles highlight the 2023 Quiché and 1997 Jutiapa earthquakes. (c) Cross section showing earthquake locations and moment-tensor orientations in a reference frame relative to the local slab surface.

tency in orientation with the rest of the intraslab deformation field. The Quiché earthquake also stands out for its magnitude – although not, by a long way, the largest intraslab event recorded within the Cocos plate, a M_w 6.4 would be expected to require a fault area on the order of ~ 100 km^2, requiring either a high-aspect-ratio rupture, or a substantial seismogenic cross section for the slab at this depth.

4 Dynamics of the Central American Slab

Simple numerical calculations in the wake of the development of early slab models explored the evolution of the intraslab stress state as a function of the slab length (e.g., Vassilou et al., 1984; Vassilou and Hager, 1988), assuming that the slab behaves as a Newtonian fluid coupled to a less viscous surrounding mantle and deforming under its own weight, descending into a layered mantle structure. Such models neglect any bending-related stresses, and any stress variations resulting from the internal rheological evolution of the slab, and therefore simply provide estimates of the intraslab stress field driven by slab pull and the interaction of the slab with the surrounding mantle. We do not attempt to reproduce the calculations of Vassilou and Hager (1988) here, but summarise their findings in Figure 6. The negative buoyancy of the slab in each case puts the shallow part of the slab into down-dip extension, whilst for slabs extending towards the mantle transition zone, where there is a significant viscosity contrast, there is a switch into down-dip compression in the deeper parts of the slab, which propagates back to increasingly shallower depths for slab that reach deeper into the mantle. The dashed line on Figure 6 shows the impact of having an inclined slab, rather than a vertical one – essentially, this serves to reduce the magnitude of the stresses involved, and produces a rotation in the local stress tensor, but has little impact on the 'polarity' of the stress field, with shallow depths still being dominated by down-dip extensional stresses. The Cocos slab under Guatemala would most closely resemble the '270' km slab model shown in Figure 6. Of note here is that the predicted stresses near the slab tip are very low, due to the small length of negatively buoyant slab extending to greater depths. We also note that for slabs only reaching to $\sim$300 km, the tip of the slab is not placed into down-dip compression, as it remains too distant from the mantle transition zone.

At this time, we are not aware of any concrete evidence that the Cocos slab persists significantly below $\sim$300 km as a contiguous structure (i.e., one which can act as a stress guide) – whilst there is clearly slab-derived material deeper into the mantle (e.g., Rogers et al., 2002), how this connects to the shallow slab is unclear, with no clear evidence for a down-dip continuous slab. In models which do image the continuation of slab-derived material below 300 km (e.g., Zhu et al., 2020; Xue et al., 2023), the weak velocity anomaly, combined with the orientation of mantle fabrics, is interpreted to show a fragmentary slab subject to through-going mantle flow. We do, however, note that further work on

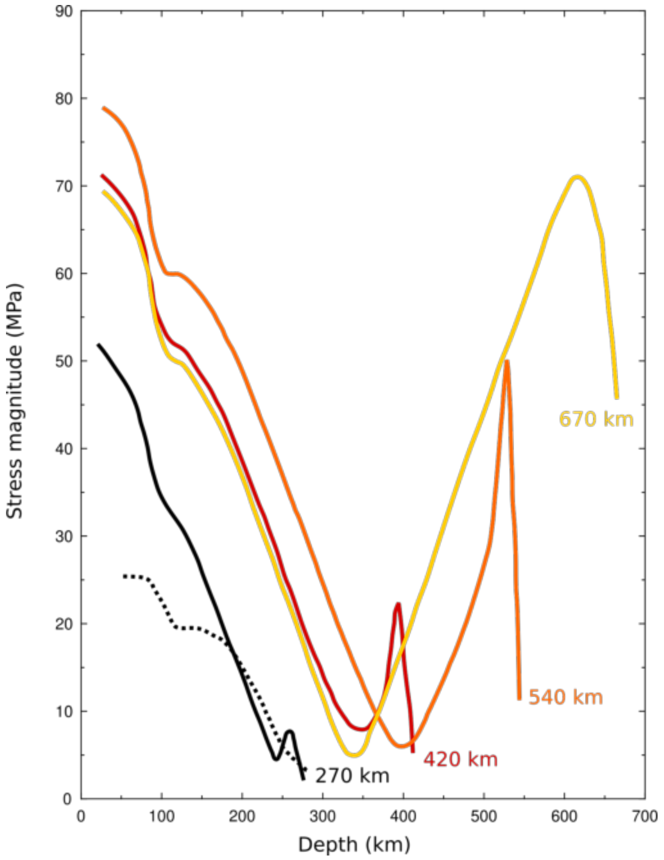


Figure 6 Stress as a function of depth for variable length slabs, descending under their own weight into a viscous mantle. All slabs are assumed to dip vertically, except for the dashed line, which dips at 45°. The in-plane stress field within the slab is in down-dip extension at shallow depths, with the increase beyond ∼350 km reflecting a switch to down-dip compression due to interaction of the slab tip with the viscosity increase at the transition zone (670 km). After Vassilou and Hager (1988).

imaging the Cocos slab beneath Central America may change this picture, and our subsequent interpretation.

Under the assumption that the Cocos slab does not persist below ∼300 km in a manner capable of sustaining significant stresses, it is notable that, even at the very tip of the contiguous Cocos slab, the slab remains capable of producing earthquakes such as those beneath Quiché and Jutiapa. The slab-pull derived intraslab stress field should, at the tip of the slab, become small, yet the consistency between the orientation of these earthquakes at 250–270 km depths, and those between 120–250 km, suggests that a similar interplay of stress field and relict structure continues to control the orientation of faulting throughout the slab, potentially modulated by the availability of a brittle rheology dependent on localised hydration along pre-existing structure. Two implications arise from this observation:

- that, in the absence of bending-related stresses, the buoyancy-related stresses (i.e., slab pull) continue to dominate the intraslab stress field, even in environments where these stresses must be small. Other sources of stress (e.g., those arising from the thermo-chemical evolution of the slab), must therefore be insignificant.
- that the slab itself must be rheologically quite weak, such that even the reduced buoyancy-related stresses present near the slab tip remain capable to activating, in the case of the Quiché earthquake, a substantial seismogenic cross section of the slab.

5 The mechanics of intermediate depth earthquakes

In keeping with other intermediate depth earthquakes (e.g., Ye et al., 2020; Wimpenny et al., 2023), both the Quiché and Jutiapa earthquakes had low-productivity aftershock sequences. Despite its own considerable magnitude (M_w 6.4), the Quiché earthquake was reported by the NEIC to be followed by only two other earthquakes within 100 km of the earthquake epicentre in the following 6 months, both considerably smaller (m_b 4.6 and 4.3) and considerably shallower ($<$ 200 km). Following the 1997 Jutiapa earthquake, the NEIC reported only one aftershock near the intermediate-depth source region, with m_b 4.9. The lack of a substantial aftershock population also inhibits the inference of earthquake rupture dimensions and causative fault plane based on aftershock distribution and extent, although we note the potential for local seismic data to clarify this (e.g., Yani-Quiyuch et al., 2023).

That the deviatoric stress derived from the negative buoyancy of the short section of the contiguous slab down-dip from these earthquakes was still capable of producing major seismogenic failure of the slab suggests that the yielding stress of the slab at such depth was also low – a condition more easily reconcilable with the rheological control on intraslab seismicity being related to either dehydration embrittlement or dehydration stress transfer, either of which would greatly reduce the effective yield stress, rather than a shear-heating model, which would still require high stresses to initiate the initial shear instability.

6 Conclusions

The Quiché and Jutiapa earthquakes represent down-dip extensional failure near the tip of the subducting Cocos slab beneath Guatemala. The coherence of the moment tensors of these earthquakes with those updip suggests that the intraslab stress field at such depths remains dominated by the negative buoyancy of the slab (slab pull). That the slab at such depths remains capable of major seismogenic failure in down-dip extension, despite the limited section of contiguous slab extending beyond the depth of these earthquakes, suggests that the slab itself is relatively weak, in order to allow failure under only low-magnitude slab pull-derived stresses. The sensitivity of the slab to relatively small deviatoric stresses (compared to lithostatic stresses), is consistent with a fluid-related rheological control on intraslab seismicity.

Acknowledgements

The authors were supported in this work by the Royal Society under URF\R1\180088 and URF\R\231019. TJC was also supported through COMET, the UK Natural Environment Research Council's Centre for the Observation and Modelling of Earthquakes, Volcanoes, and Tectonics. The numerical modelling component of this work was undertaken on ARC4, part of the High Performance Computing facilities at the University of Leeds, UK.

Data and code availability

We are extremely grateful to all those involved in the deployment and maintenance of the seismological networks whose data we draw on. We draw on seismological catalogues from the gCMT project (https://www.globalcmt.org; last accessed 22/1/2023), and from the International Seismological Centre (http://www.isc.ac.uk; last accessed 22/1/2023). This work principally used code provided as part of the Generic Mapping Tools (Wessel et al., 2013), and via the Pyrocko toolbox (https://pyrocko.org, Heimann et al., 2017). Waveform inversions were carried out using the Grond software of Heimann et al. (2018).

Competing interests

The authors declare no known competing interests.

References

Bailey, I. W., Alpert, L. A., Becker, T. W., and Miller, M. S. Co-seismic deformation of deep slabs based on summed CMT data. *Journal of Geophysical Reearch*, 117, 2012. doi: 10.1029/2011JB008943.

Boneh, Y., Schottenfels, E., Kwong, K., van Zelst, I., Tong, X., Eimer, M., Miller, M., Moresi, L., Warren, J., Wiens, D., nd J. Naliboff, M. B., and Zhan, Z. Intermediate-depth earthquakes controlled by incoming plate hydration along bending-related faults. *Geophysical Research Letters*, 46:3688–3697, 2019. doi: 10.1029/2018GL081585.

Craig, T., Methley, P., and Sandiford, D. Imbalanced moment release within subducting places during initial bending and unbending. *Journal of Geophysical Research*, 127, 2022. doi: 10.1029/2021JB023658.

Hayes, G., Moore, G., Portner, D., Hearne, M., Flamme, H., Furtney, M., and Smocyzk, G. Slab2, a comprehensive subduction zone geometry. *Science*, 2018. doi: 10.1126/science.aat4723.

Heimann, S., Kriegerowski, M., Isken, M., Cesca, S., Daout, S., Grigoli, F., Juretzek, C., Megies, T., Nooshiri, N., Steinberg, A., Sudhaus, H., Vasyura-Bathke, H., Willey, T., and Dahm, T. Pyrocko - An open-source seismology toolbox and library. *GFZ Data Services*, 2017. doi: 10.5880/GFZ.2.1.2017.001.

Heimann, S., Isken, M., Kuhn, D., Sudhaus, H., Steinberg, A., Vasyura-Bathke, H., Daout, S., Cesca, S., and Dahm, T. Grond - A probabilistic earthquake source inversion framework. *GFZ Data Services*, 2018. doi: 10.5880/GFZ.2.1.2018.003.

Hosseinzadehsabeti, E., Ferre, E., Persaud, P., Fabbri, O., and Geissman, J. The rupture mechanisms of intraslab earthquakes: a multiscale review and re-evaluation. *Earth Science Reviews*, 221, 2021. doi: 10.1016/j.earscirev.2021.103782.

Isacks, B. and Molnar, P. Mantle earthquake mechanisms and the sinking of the lithosphere. *Nature*, 223:1121–1124, 1969.

Kennett, B. L. N., Engdahl, E. R., and Buland, R. Constraints on seismic velocities in the Earth from travel times. *Geophysical Journal International*, 122:108–124, 1995.

Kim, Y., Clayton, R., and Jackson, J. Geometry and seismic properties of the subducting Cocos plate in central Mexico. *Journal of Geophysical Research*, 115, 2010. doi: 10.1029/2009JB006942.

Manea, V., Manea, M., Kostoglodov, V., and Sewell, G. Intraslab seismicity and thermal stress in the subducted Cocos plate beneath central Mexico. *Tectonophysics*, 420, 2006. doi: 10.1016/j.tecto.2006.03.029.

Manea, V., Manea, M., and Ferrari, L. A geodynamical perspective on the subduction of Cocos and Rivera plates beneath Mexico and Central America. *Tectonophysics*, 609:56–81, 2013. doi: 10.1016/j.tecto.2012.12.039.

Manea, V., Manea, M., Ferrari, L., Orozco-Esquivel, T., Valenzuela, R., Husker, A., and Kostoglodov, V. A review of the geodynamic evolution of flat slab subduction in Mexico, Peru and Chile. *Tectonophysics*, 695:27–52, 2017. doi: 10.1016/j.tecto.2016.11.037.

Masson, D. G. Fault Patterns at Outer Trench Walls. *Marine Geophysical Researches*, 1991:209–225, 1991.

Melgar, D., Pérez-Campos, X., Ramirez-Guzman, L., Spica, Z., Espíndola, V., Hammond, W., and Cabral-Cano, E. Bend Faulting at the Edge of a Flat Slab: The 2017 M_W 7.1 Puebla-Morelos, Mexico, Earthquake. *Geophysical Research Letters*, 45:2633–2641, 2018a. doi: 10.1002/2017GL076895.

Melgar, D., Ruiz-Angula, A., Garcia, E., Manea, M., Manea, V., Xu, X., Ramirez-Herrera, M., Zavala-Hidalgo, J., Geng, J., Corona, N., Pérez-Campos, X., Cabral-Cano, E., and Ramirez-Guzmán, L. Deep embrittlement and complete rupture of the lithosphere during he M_w 8.2 Tehuantepec earthquake. *Nature Geoscience*, pages 955–960, 2018b. doi: 10.1038/s41561-018-0229y.

Prakash, A., Holyoke III, C., Kelemen, P., Kirby, S., Kronenberg, A., and Lamb, W. Carbonates and intermediate-depth seismicity: Stable and unstable shear in altered subducting plates and overlying mantle. *Proceedings of the National Academy of Sciences*, 120, 2023. doi: 10.1073/pnas.2219076120.

Ranero, C. R., Villaseñor, A., Morgan, J. P., and Weinrebe, W. Relationship between bend-faulting at trenches and intermediate-depth seismicity. *Geochemistry, Geophysics, Geosystems*, 6, 2005. doi: 10.1029/2005GC0009972.

Rogers, R., Kárason, H., and van der Hilst, R. Epeirogenic uplift above a detached slab in northern Central America. *Geology*, 30:1031–1034, 2002. doi: 10.1130/0091-7613(2002)030.

Sandiford, D., Moresi, L., Sandiford, M., and Yang, T. Geometric controls on flat slab seismicity. *Earth and Planetary Science Letters*, 527, 2019. doi: 10.1016/j.epsl.2019.115787.

Sandiford, D., Moresi, L., Sandiford, M., Farrington, R., and Yang, T. The Fingerprints of Flexure in Slab Seismicity. *Tectonics*, 39, 2020. doi: 10.1029/2019TC005894.

Singh, S., Ordaz, M., Alcántara, L., Shapiro, N., Kostoglodov, V., Pacheco, J., Alcocer, S., Gutiérrez, C., Quass, R., Mikumo, T., and Ovando, E. The Oaxaca Earthquake of 30 September 1999 (Mw = 7.5): A Normal-faulting Event in the Subducted Cocos Plate. *Seismological Research Letters*, 71:67–78, 2000. doi: 10.1785/gssrl.71.1.67.

Sippl, C., Dielforder, A., John, T., and Schmalholz, S. Global Constraints on Intermediate-Depth Intraslab Stresses From Slab Geometries and Mechanisms of Double Seismic Zone Earthquakes. *Geochemistry Geophyics Geosystems*, 23, 2022. doi: 10.1029/2022GC010498.

Syracuse, E., Abers, G., Fischer, K., MacKenzie, L., Rychert, C., Protti, M., González, V., and Strauch, W. Seismic tomography and earthquake locations int he Nicaragua and Costa Rican upper mantle. *Geochemistry Geophysics Geosystems*, 9, 2008. doi: 10.1029/2008GC001963.

Vallée, M., Bouchon, M., and Schwartz, S. Y. The 13 January 2001 El Salvador earthquake: A multidata analysis. *Journal of Geophysical Research*, 108, 2003. doi: 10.1029/2002JB001922.

Vassilou, M. and Hager, B. Subduction Zone Earthquakes and Stress in Slabs. *Pure and Applied Geopysics*, 128:547–624, 1988. doi: 10.1007/BF00874550.

Vassilou, M., Hager, B., and Raefsky, A. the distribution of earthquakes with depth and stress in subducting slabs. *Journal of Geodynamics*, 1:11–28, 1984. doi: 10.1016/0264-3707(84)90004-8.

Wessel, P., Smith, W., Scharroo, R., Luis, J., and Wobbe, F. Generic Mapping Tools: Improved version released. *EOS Transactions*, 94:409–410, 2013.

Wimpenny, S., Craig, T., and Marcou, S. Re-Examining Temporal Variations in Intermediate-Depth Seismicity. *Journal of Geophysical Research*, 128, 2023. doi: 10.1029/2022JB026269.

Xue, T., Peng, D., Liu, K., Obrist-Farner, J., Locmelis, M., Gao, S., and Liu, L. Ongoing fragmentation of the usubducting Cocos slab, Central America. *Geology*, 51, 2023. doi: 10.1130/G51403.1.

Yani-Quiyuch, R., Asturias, L., and Castro, D. The rupture plane of the 16 February 2022 M_w 6.2 Guatemala, intermediate depth earthquake. *Seismica*, 2.2, 2023. doi: 10.26443/seismica.v2i2.691.

Ye, L., Lay, T., and Kanamori, H. Anomalously low aftershock productivity of the 2019 Mw 8.0 energetic intermediate-depth faulting beneath Peru. *Earth and Planetary Science Letters*, 549, 2020. doi: 10.1016/j.epsl.2020.116528.

Zhu, H., Stern, R., and Yang, J. Seismic evidence for subduction-induced mantle flows underneath Middle America. *Nature Communications*, 11, 2020. doi: 10.1038/s41467-020-15492-6.

SEISMICA

Mapping fault geomorphology with drone-based lidar

Guy Salomon * [1], **Theron Finley** [1], **Edwin Nissen** [1], **Roger Stephen** [1,2], **Brian Menounos** [3,4]

[1]School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia, Canada, [2]Department of Geography, University of Victoria, Victoria, British Columbia, Canada, [3]Department of Geography, Earth, and Environmental Sciences, University of Northern British Columbia, Prince George, British Columbia, Canada, [4]Geological Survey of Canada, PO Box 6000, Sidney, BC V8L 4B2, Canada

Author contributions: *Conceptualization*: E.N., G.S., T.F. *Methodology*: G.S., R.S. *Formal Analysis*: G.S. *Investigation*: G.S., T.F., R.S., E.N., B.M. *Writing - original draft*: G.S.. *Writing - Review & Editing*: E.N., T.F, B.M. *Funding acquisition*: E.N., T.F., B.M.

Abstract The advent of sub-meter resolution topographic surveying has revolutionized active fault mapping. Light detection and ranging (lidar) data collected using crewed airborne laser scanning (ALS) can provide ground coverage of entire fault systems but is expensive, while Structure-from-Motion (SfM) photogrammetry from uncrewed aerial vehicles (UAVs) is popular for mapping smaller sites but cannot image beneath vegetation. Here, we present a new UAV laser scanning (ULS) system that overcomes these limitations to survey fault-related topography cost-effectively, at desirable spatial resolutions, and even beneath dense vegetation. In describing our system, data acquisition and processing workflows, we provide a practical guide for other researchers interested in developing their own ULS capabilities. We showcase ULS data collected over faults from a variety of terrain and vegetation types across the Canadian Cordillera and compare them to conventional ALS and SfM data. Due to the lower, slower UAV flights, ULS offers improved ground return density (∼260 points/m^2 for the capture of a paleoseismic trenching site and ∼10–72 points/m^2 for larger, multi-kilometer fault surveys) over conventional ALS (∼3–9 points/m^2) as well as better vegetation penetration than both ALS and SfM. The resulting ∼20–50 cm-resolution ULS terrain models reveal fine-scale tectonic landforms that would otherwise be challenging to image.

Production Editor:
Gareth Funning
Handling Editor:
Randolph Williams
Copy & Layout Editor:
Oliver Lamb

Received:
January 18, 2024
Accepted:
May 5, 2024
Published:
June 3, 2024

Non-technical summary Lidar remote sensing uses light pulses from a laser instrument to measure distances to objects and surfaces to create high precision, three-dimensional models. It is useful for mapping the ground surface where obscured by forest, because laser pulses that avoid foliage and branches will sample the ground surface while those that don't can be digitally removed, unlike in photogrammetry. Typically, lidar instruments are mounted on tripods (terrestrial laser scanning) or crewed aircraft (airborne laser scanning). Recently, lidar systems have become compact and light enough to be deployed from uncrewed aerial vehicles (UAVs, or drones) and this technology is being adopted across many disciplines. Here, we describe some of the first applications of a drone lidar system to study landforms generated by active faults, and illustrate its capabilities using surveys of a variety of faulted landscapes with different vegetation types across western Canada. Our system offers a cost-effective way of obtaining otherwise expensive lidar data, and compares favourably against established methods of topographic mapping, allowing us to survey the landscape in finer detail than was previously possible. The drone system is subject to practical and regulatory constraints and we discuss ways that these could be mitigated in the future.

1 Introduction

Lidar (light detection and ranging) is an increasingly popular terrestrial remote sensing method that combines the return times of reflected or back-scattered laser pulses with information on the location and orientation of the laser scanner to produce a dense 'point cloud' containing the Cartesian (x, y and z) co-ordinates of a geographic target (Xiaoye Liu, 2008; Glennie et al., 2013). The sub-meter point spacings characteristic of lidar data are finer than the ∼1–10 meter pixel dimensions typical of modern satellite-derived digital elevation models (DEMs) (e.g. Morin et al., 2016; Hodge et al., 2019; Wang et al., 2019; Benavente et al., 2021; Salomon et al., 2022). Furthermore, since multiple laser returns can be distinguished from the same outgoing pulse, and since distinct canopy returns can be filtered out, lidar is able to penetrate vegetation to yield a bare-earth digital terrain model (DTM) of the ground surface. These unique attributes of lidar remote sensing have contributed to an explosion of interest across many geospatial fields, including tectonic geomorphology (Meigs, 2013). It is becoming common practice to acquire lidar along fault surface traces as it provides some of the best data for constraining fault offsets, kinematics, and scarp morphology (e.g. Cunningham et al., 2006; Hilley et al., 2010; Zielke et al., 2010, 2015; Elliott et al., 2012; Salisbury et al., 2012; Johnson et al., 2018; Wei et al., 2019), as well as a topographic baseline for mapping any future earthquake deformation (Oskin et al., 2012; Glennie et al., 2014; Nissen et al., 2014; Scott et al., 2018; Diederichs et al., 2019; Ishimura et al., 2019; Lajoie et al., 2019; Wedmore et al., 2019) or aseismic fault

*Corresponding author: guysalomon@uvic.ca

creep (DeLong et al., 2015; Scott et al., 2020). Lidar is especially useful in regions such as northwestern North America, where widespread forest cover may otherwise obscure fault scarps or other earthquake-related landforms (e.g. Haugerud et al., 2003; Hunter et al., 2011; Morell et al., 2017; Nelson et al., 2017; Johnson et al., 2018; Harrichhausen et al., 2021; Schermer et al., 2021; Witter et al., 2021).

Lidar data are typically collected through one of two established methods. In Terrestrial Laser Scanning (TLS), landscapes are mapped at low incidence angles from a laser instrument mounted on a stationary tripod (Telling et al., 2017). TLS can achieve very dense point clouds with 100s to 1000s of points per square meter (pts/m^2), but to avoid shadowing of features of interest behind objects like tree trunks or undulating topography, the scanner is typically deployed at several locations. Furthermore, due to the time taken to set up each new scanner position, TLS surveys are best suited for relatively small outcrop or landform-scale acquisitions (e.g. Jones et al., 2009; Haddad et al., 2012; Gold et al., 2013; Wiatr et al., 2013; Bubeck et al., 2015; DeLong et al., 2015; Wedmore et al., 2019). In rare instances, terrestrial lidar surveys have been expanded by mounting the scanner on motorized vehicles, backpacks, or tethered balloons, a configuration termed Mobile Laser Scanning (MLS) (Glennie et al., 2013; Brooks et al., 2013; Brooks et al., 2017; Nevitt et al., 2020; Zhu et al., 2022). The second, more prevalent method is Airborne Laser Scanning (ALS), where the laser scanner is mounted onto a crewed aircraft and flown over the target area (Xiaoye Liu, 2008; Glennie et al., 2013). This method is suitable for collecting much larger datasets, typically in 1–2 km wide swaths that can extend tens to hundreds of kilometers along linear targets such as fault traces, coastlines, or infrastructure corridors (e.g. Toth et al., 2007; Prentice et al., 2009; Hubbard et al., 2011; Oskin et al., 2012; Langridge et al., 2014; Clark et al., 2017; Johnson et al., 2018). ALS generally yields much lower point densities than TLS with typical values for modern acquisitions being 10–15 pts/m^2. Other limitations associated with ALS include the steep cost of deploying a crewed fixed-wing aircraft (10s to 100s of thousands of dollars per survey), restrictions on what altitudes and speeds they can fly at (which limits the raw point density), and constraints on scanning angles that can prevent penetration of the densest vegetation (VanValkenburgh et al., 2020; Resop et al., 2019). Rotary-wing aircraft (helicopters) are less constrained in terms of altitude and speed and have been used to achieve higher point density in some lidar surveys (Chen et al., 2015), but are generally even more cost-prohibitive.

Recently, a proliferation of inexpensive uncrewed aerial vehicles (UAVs)—commonly referred to as ‘drones’ and formally as remotely piloted aircraft systems (RPAS)—have provided a more accessible means of terrain mapping, including for seismology and active tectonics (e.g. Bemis et al., 2014; Johnson et al., 2014; DuRoss et al., 2019). Until very recently, this has involved deploying cameras and using sophisticated photogrammetric algorithms to create Digital Surface Models (DSMs) (e.g. Harwin and Lucieer, 2012; James and Robson, 2012; Westoby et al., 2012), with consequently only very limited ability to map forested landscapes. However, the recent development of smaller, lighter laser scanners has opened up the possibility of collecting lidar datasets from UAV platforms, referred to from now on as drone lidar or UAV laser scanning (ULS) (Wieser et al., 2016). This new advancement has seen some early adoption in the fields of archaeology (Risbøl and Gustavsen, 2018; VanValkenburgh et al., 2020), forestry and ecology (Brede et al., 2017; Kellner et al., 2019; Tomsett and Leyland, 2021), and fluvial and landslide geomorphology (Resop et al., 2019; Pellicani et al., 2019), but its effectiveness for mapping active faulting has not yet been demonstrated.

This paper introduces a state-of-the-art ULS system (Fig. 1) developed at the University of Victoria to study the geomorphology of putative active faults across the Canadian Cordillera. This is a region of widespread seismicity (Ristau et al., 2007) and elevated seismic hazard (Kolaj et al., 2020), but aside from the major plate boundary faults, only a few seismogenic faults have been conclusively identified and mapped owing to steep terrain, dense forest cover, and recent glaciation (e.g. Morell et al., 2017; Harrichhausen et al., 2023). We begin in Section 2 by describing the drone platform and instrumentation as well as our data collection, processing and analysis workflows. Our aim here is to provide a blueprint for other seismologists and geomorphologists interested in developing their ULS systems. In Sections 3–6, we then showcase examples of ULS lidar data collected using our drone platform along four faults with differing surface expressions in four unique types of vegetation cover from across western Canada (Fig. 2 and Table 2). The spatial scales of these case studies vary from a paleoseismic trench site with dimensions of a few hundred meters surveyed in a few hours (Harrichhausen et al., 2023) to regional acquisitions along fault sections totalling several kilometers in length mapped over several days (Finley et al., 2022a). In each case study, we compare the ULS data both quantitatively and qualitatively with existing ALS data, as well as SfM data where available. Particular focus will be on vegetation penetration performance, achievable ground return densities, and their impacts on derived DTM resolutions for the different lidar acquisition modes. In Section 7, we assess the impact of UAV flight speed on survey duration, point density and data quality, before discussing the unique applications, advantages, and limitations of ULS in active tectonics research. One of the principal advantages is simplified and cheaper repeat observations, with the potential for imaging co-seismic rupture, off-fault deformation, and post-seismic afterslip at finer spatial (<50 cm) and temporal (<1 day) resolutions. Additionally, the higher spatial resolutions attainable with ULS allow for more confident measurement and interpretation of subtle fault scarp morphology. The limitations of ULS systems include spatial coverage, which is restricted by battery life, flight constraints imposed by civil aviation authorities, and the

Term	Description
AGL	Above ground level.
ALS	Airborne laser scanning—lidar from a crewed aircraft. Synonymous with ALSM and airborne lidar.
ALSM	Airborne laser swath mapping—lidar from a crewed aircraft. Synonymous with ALS and airborne lidar.
BVLOS	Beyond visual line-of-sight.
DEM	Digital Elevation Model—a 3-D representation of terrain heights. Synonymous with DTM.
DoD	DEM of Difference—an elevation difference map between two DEMs.
DSM	Digital Surface Model—a 3-D representation of Earth surface heights, incl. natural or man-made objects.
DTM	Digital Terrain Model—a 3-D representation of terrain heights. Synonymous with DEM.
GCP	Ground control point—an identifiable point on Earth's surface with known location used for georeferencing.
GNSS	Global Navigation Satellite Systems—a system that uses satellites to provide autonomous geopositioning.
GPS	Global Positioning Systems—the world's most utilized GNSS, and sometimes used synonymously.
ICP	Iterative closest point—an algorithm used to co-register two point clouds.
IMU	Inertial Measurement Unit—a device that tracks orientation using magnetometer, accelerometer, and gyro.
INS	Inertial Navigation System—a device integrating an IMU and GNSS.
.LAS	Industry standard binary file format used for the interchange and archiving of lidar data.
.LAZ	Compressed file format used for the interchange and archiving of lidar data.
lidar	Light detection and ranging, with varied capitalization (LiDAR, LIDAR, Lidar). Synonymous with laser scanning.
MLS	Mobile laser scanning—lidar from a roving scanner on or tethered to Earth's surface.
M3C2	Multiscale Model to Model Cloud Comparison— a method that measures differences between point clouds.
PPP	Precise Point Positioning—a GNSS method that calculates positions with errors of a few centimeters.
RINEX	Receiver Independent Exchange Format—a data interchange format for raw GNSS data.
RPAS	Remotely Piloted Aircraft System. Synonymous with UAV and drone.
RTH	Return to Home—a feature of some UAS that allows the drone to return autonomously to its take-off point.
sbet	Smoothed best estimate of trajectory—relating to the processing of UAV flight paths.
SfM	Structure-from-Motion—an algorithm for estimating 3-D scene structure from a set of photographs.
sUAS	Small Unmanned Aircraft System—a UAV weighing less than 25 kg.
TLS	Terrestrial Laser Scanning—lidar from one or more stationary locations on Earth's surface.
UAS	Uncrewed Aircraft System—a UAV and its accessories (e.g. ground control, transmission).
UAV	Uncrewed Aerial Vehicle. Synonymous with RPAS or drone.
ULS	UAV Laser Scanning—lidar from a UAV platform, also referred to here as "drone lidar".
VLOS	Visual line-of-sight.

Table 1 Acronyms and initialisms used in this paper and/or common within the wider literature on drone-based and lidar remote sensing, with abbreviated definitions where helpful.

Section	Fault	Target and landscape description	Vegetation Type	Area (km^2)	Time to Collect
3	X̲EOLX̲ELEK̶ - Elk Lake fault, Vancouver Island	Reverse fault scarp within a suburban park	Pacific Dry Forest	0.01	5 hours (1 day)
4	San Juan fault, Cowichan Valley, Vancouver Island	Strike-slip fault within a steep-sided valley	Pacific Cool Temperate Forest	2.8	14 hours (2 days)
5	Southern Rocky Mountain Trench fault, East Kootenay	Suspected fault scarp crossing gently-sloping alluvial fans	Cordilleran Dry Forest	3.1	16 hours (3 days)
6	Eastern Denali fault, Kluane Lake	Major strike-slip fault in a broad glacial valley	Northwestern Boreal forest	10.4	12 days

Table 2 Summary of case studies described in this paper. The drone lidar platform was tested in several different tectonic settings and climatic regions along faults of a variety of kinematic styles within the Canadian Cordillera in British Columbia and the Yukon. Canadian vegetation types are from Baldwin et al. (2019).

necessity of road access to launch sites with good visual sightlines, and we finish the paper by discussing ways in which these limitations might be mitigated in the future.

2 Methods

2.1 The ULS system

Our ULS platform was built using several commercially available and custom-built components (Fig. 1). The former comprise a DJI Matrice 600 Pro hexacopter, a Riegl miniVUX-1UAV laser scanner, an Applanix APX-20 UAV Inertial Navigation System (INS), and a Trimble AV14 antenna. Custom-built elements include an interface board used to integrate the laser scanner and INS and a housing and mounting mechanism. The 2.75 kg payload is mounted to the drone with a dovetail-style connector, similar to those used in motion-stabilized gimbals for cinematography.

The DJI Matrice 600 Pro hexacopter has a payload capacity of 6 kg, more than double what we deploy. It uses one set of six TB47S Lithium-Ion Polymer batteries per flight (4500 mAh), enough power for ∼20 minutes of flying with our payload. Notably, these batteries are just below the 100 watt-hour rating restriction imposed by civil aviation authorities, allowing us to travel to field sites with the drone system via commercial airline. The drone is maneuvered via a remote controller and transmitter with a manufacturer-stated maximum operating distance of 5 km, though in practice, we begin to encounter connectivity issues beyond ∼1.5 km in forested and mountainous terrain. The pilot must abide by flight constraints imposed by civil aviation authorities, including altitude limits of 400 ft (121.92 m) above ground level (AGL) in both Canada and the U.S., restrictions to flights over people, and maintenance of visual line-of-sight between the pilot (or a visual observer in constant radio contact with the pilot) and the drone. Many countries have similar UAV regulations, although exact parameters for flight height and horizontal distance do vary (Stöcker et al., 2017). In Canada, pilots must have obtained Advanced Operations drone pilot certification from Transport Canada (Transport Canada, 2022) through an online multiple-choice test and an in-person flight review. This certification allows the pilot to operate in controlled airspace with any drone weighing less than 25 kg. Many other national aviation authorities offer similar certifications (e.g. Federal Aviation Administration, 2023; UK Civil Aviation Authority, 2023; European Union Aviation Safety Authority, 2022; Civil Aviation Safety Authority, 2021).

Weighing 2 kg and with a manufacturer-stated optimum altitude of 80 m AGL, the miniVUX-1UAV laser scanner is specifically designed for deployment from a drone. It offers an eye-safe laser (at Laser Class 1) with a pulse repetition rate of 100 kHz and a 360° field of view. The laser footprint diameter at optimum altitude is 6.4 cm at nadir and 9 cm at 45° from nadir. The scanner can record up to 5 returns from a single laser pulse, making it suitable for application in densely vegetated areas where SfM terrain mapping would be unfeasible. The Applanix APX-20 UAV INS integrates a Global Navigation Satellite Systems (GNSS) device and an Inertial Measurement Unit (IMU), which together with the attached Trimble AV14 antenna track the precise location and orientation of the laser scanner. This allows the coordinates of points within the final point cloud to have sub-decimetric accuracies (3–5 cm). The breakout board interfaces with the laser scanner using a customized circuit board which allows communication and timing between components, streams data between the INS and the laser scanner, and distributes power to both of these systems.

Our ULS system also makes use of a range of auxiliary equipment (Figure 3A). This includes a Trimble R12 GNSS base station (and tripod) for post-processing the drone trajectory and—if the best possible absolute georeferencing is desired—a separate Trimble R12 GNSS rover unit, TSC7 handheld computer and monopod for surveying ground control points (GCPs). For these we use 120 cm × 120 cm fabric harlequin-iron cross targets, which we secure to the ground with hammer and nails. We also pack a field laptop with flight planning software installed, an iPad or cell phone to connect to the radio controller, walkie-talkies to allow constant communication between crew members, an inverter generator (minimum 2200 watts to support all charging needs), charging equipment and spare drone batteries to allow a quick succession of repeat flights, and field safety gear. All of the equipment fits inside our Jeep Wrangler field vehicle with room for three crew members and their personal gear.

2.2 Survey planning

Initial planning starts with consideration of three factors. Firstly, as our drone platform and auxiliary equipment (Figure 3) are too bulky to be carried easily by foot, launch sites must be accessible via vehicle. Existing coarse resolution DTMs, satellite imagery, and Google Streetview are great tools for identifying such spots. In areas with steep topography or forest cover, visual-line-of-sight is often impossible to maintain from the launch site and so we use one or more visual observers positioned in areas with good sight lines in constant communication with the drone pilot via walkie-talkie. Secondly, if the survey is located within classified airspace, approval must be applied for in advance. In Canada this can be done online through NAV Canada's NAV drone application. Thirdly, in the days leading up to the fieldwork we check the weather forecasts, as our drone cannot operate in any form of precipitation or in winds that are greater than 8 m s^{-1}.

Once these initial considerations are addressed, we use drone flight planning software to generate automated flight paths for our data collection. Map Pilot Pro by Maps Made Easy and Universal Ground Control Station (UgCS) by SPH Engineering have both been used successfully for this purpose, with the latter our current preference since it allows more control over survey design (custom base maps and images), flight parameters (e.g., altitude, speed), and laser parameters (e.g., field of view, swath overlap) such that a desired point den-

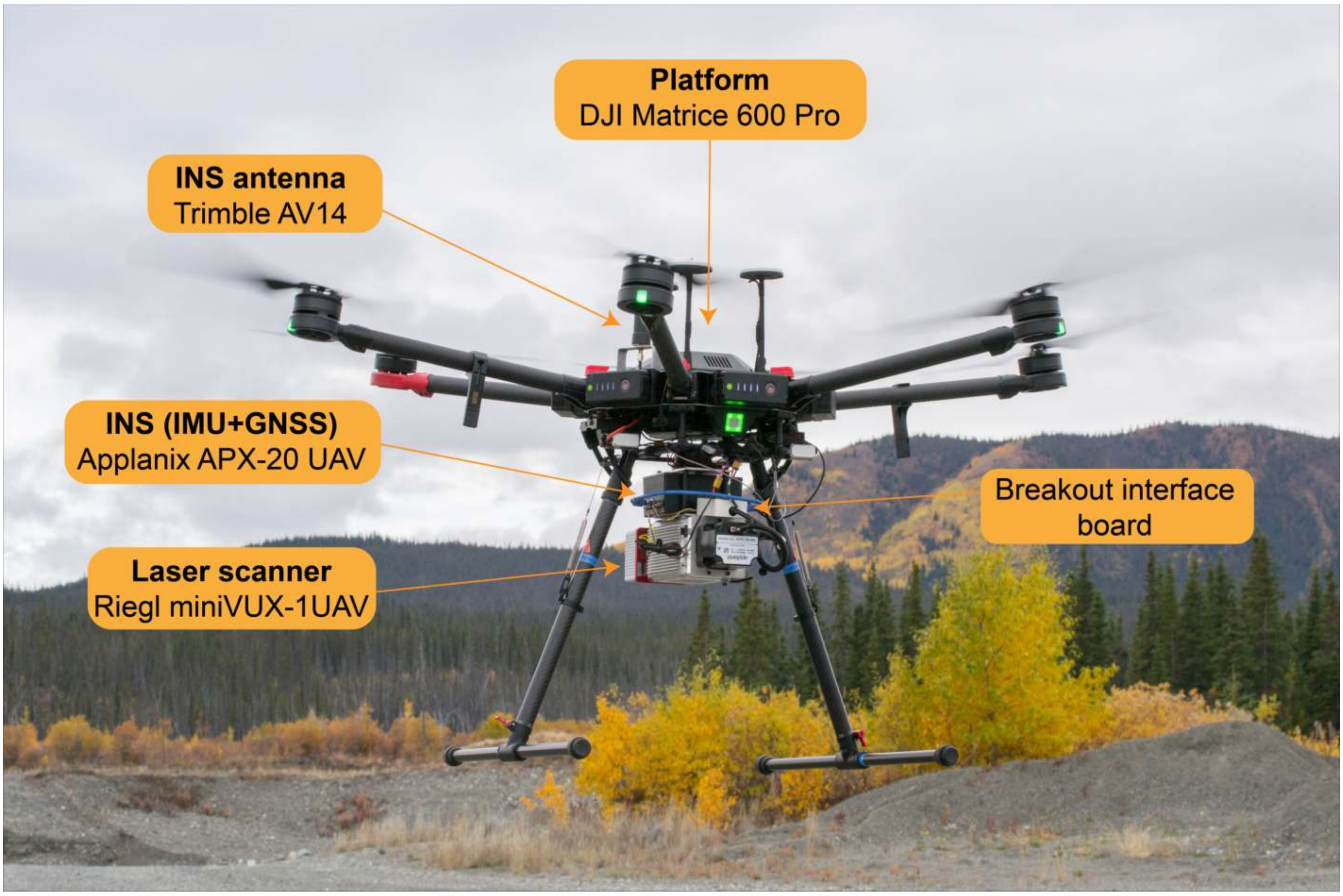


Figure 1 Annotated photograph of the drone platform and instrumentation used in this study. For scale, the full diameter of the drone, including rotor blades, is 1.66 m.

sity is achieved. There is a trade off between point density and areal coverage and the specific scientific goals of the survey need to be considered. These parameters are fine-tuned using Riegl's RiParameter software to ensure even point spacing, and set in the laser instrument using the RiAcquire tool. We discuss optimization of these parameters in section 7.1. Typical flight plans for fault-related studies consist of 2–8 strike parallel survey lines, and 2 cross-track lines for the purposes of track alignment in post-processing (Figure 3B). For larger surveys these will be undertaken in multiple flights in order to allow for battery replacement. Note that each individual flight must have a minimum of two overlapping lines, and it is best practice to design surveys that allow for the completion of full lines, rather than abandoning and resuming part-way along a line; overlapping data is critical for scanline alignment when merging flights in post-processing. In fault surveys when the desire is to achieve maximum coverage along-strike, the most efficient flight plans in our experience consume ~35% battery on the outward track and ~35% on the return track, allowing the drone to return to home safely at 30% battery, the depletion threshold recommended by the manufacturers. Survey extents are further limited by the need to maintain visual line-of-sight between the pilot, or one or more visual observers in constant radio contact with the pilot, and the drone. In the absence of obstacles we find this visual limit to be around ~1.5 km, though it is often challenging in forested areas to find ideal sight lines. In practice, considering battery, radio controller connectivity, and line-of-sight requirements, we find that survey lengths from an individual launch site are limited to a maximum of ~1.8 km even with visual observers present. Flight paths for the case studies presented in this study are provided in the supplemental material (SM1–11) to illustrate survey patterns required for different spatial extents and terrain.

2.3 Data acquisition

On the day of the survey, we drive to our launch site where we first set up the GNSS base station (Figure 3C). Getting the base recording started early ensures that there is a sufficiently long (minimum 3 hour) base observation for post processing the flight trajectories. The drone platform is assembled and the laser and IMU payload mounted onto it. Before uploading the automated flight plan, a short, manual test flight ensures that the controller is operating as expected and that the drone's gyroscope and magnetometer are calibrated. If desired, GCPs can be placed in 4–6 locations scattered across the survey footprint and clear of forest cover. These locations should ideally be situated beneath multiple overlapping and orthogonal flight lines. It is expedient to deploy GCPs before or during the drone setup and test flight, often by our visual observers as they move into position. In order to calibrate the IMU, the drone system is powered up and sits for 5 minutes of static cali-

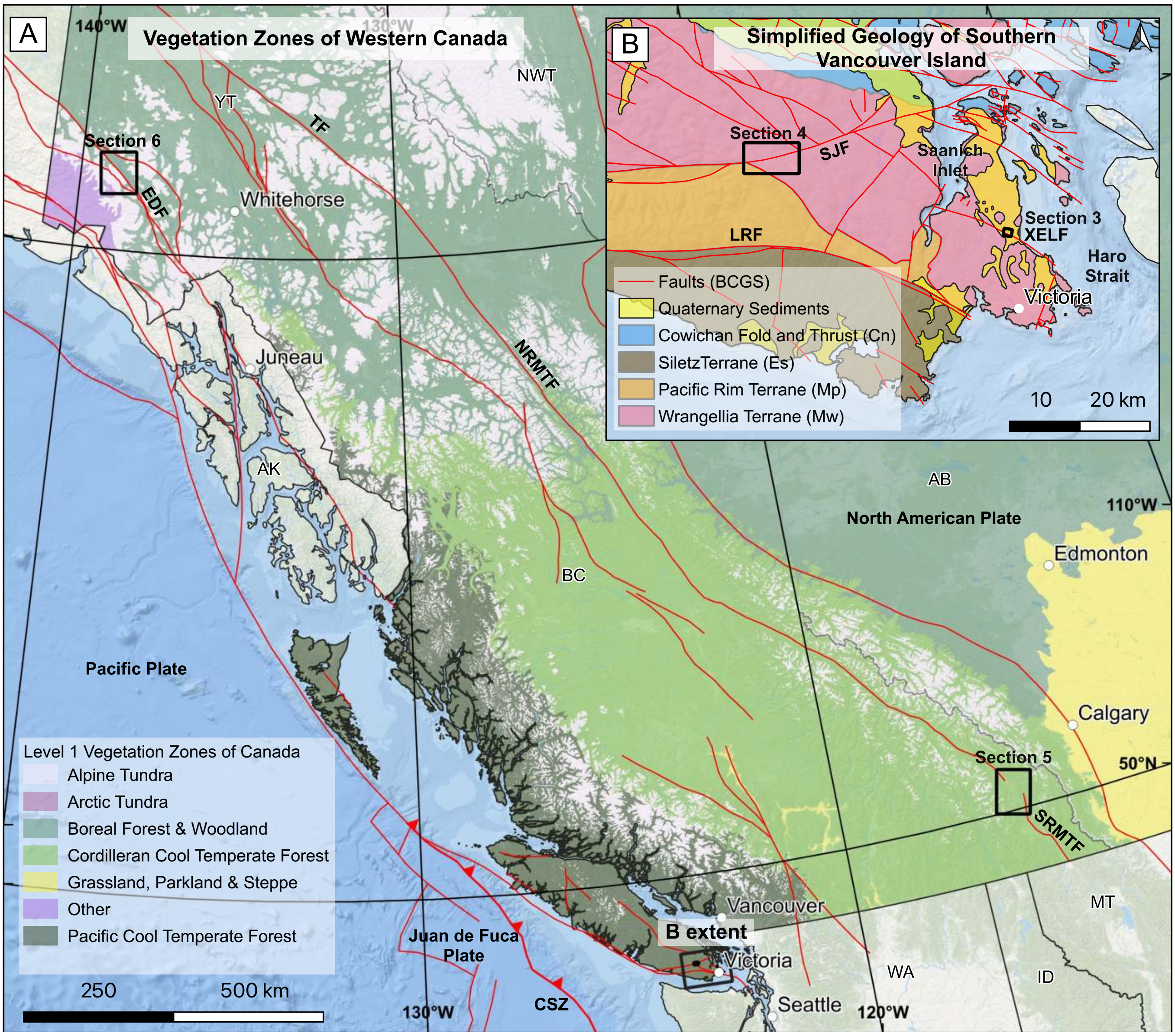


Figure 2 Location of case study sites (Table 2). (A) Level 1 Vegetation zones for Canada from Baldwin et al. (2019). Major faults are in red, some of which have been labelled. CSZ: Cascadia Subduction Zone, EDF: Eastern Denali fault, NRMTF: Northern Rocky Mountain Trench fault, SRMTF: Southern Rocky Mountain Trench fault, TF: Tintina fault. (B) Simplified geological map for southern Vancouver Island with major faults labelled, modified from the BC Geological Survey compilation by Cui et al. (2017). LRF: Leech River fault, SJF: San Juan fault, XELF: X̱EOLX̱ELEḴ-Elk Lake fault. Abbreviations for states, provinces and territories are as follows; AB: Alberta, AK: Alaska, BC: British Columbia, ID: Idaho, MT: Montana, NWT: Northwest Territories, WA: Washington, YT: Yukon. Background imagery is from Esri (2022).

bration, followed by a dynamic calibration that involves accelerating, decelerating and strafing to the left and right. Finally, the flight plan is uploaded to the drone from our field laptop or tablet, after making any last minute adjustments as needed. The entire set-up period, from arriving at the launch site to the start of the first survey flights, typically takes an hour with two or three people present.

The drone is then launched and its automated flight pattern started, with pilot and visual observers in constant radio communication to ensure that it is always within sight and maintaining sufficient clearance of obstacles. Once the batteries approach the 30% depletion threshold, the drone is brought back to the landing spot, and the static calibration is repeated before the system can be powered down and the batteries changed for fresh ones. For larger surveys, it is necessary to bring several sets of batteries and generator to recharge depleted batteries and keep flying throughout the day. At our highest levels of operating efficiency, we find that 6 sets of 6 batteries (TB47S) and two DJI Hex Chargers, running simultaneously and continuously, are necessary to keep pace with surveying. Once the full survey has been successfully flown, the base station needs to remain running for a minimum of half an hour to ensure that its location is well constrained, as is recommended in both Riegl and Applanix documentation. If we are surveying GCPs with a GNSS rover, we usually do this after the final drone flight, and leave the base station running yet another half an hour as we pack up the remaining gear.

2.4 Data Processing

After a successful survey, data from the laser scanner, the INS, and the GNSS base station and rover are copied to a workstation for processing. We follow the workflow summarized in Figure 4, which includes several pre-processing steps before the final point cloud is generated. The first step involves processing the GNSS data collected by the base station and rover. The base observation file is converted to RINEX format and uploaded to Natural Resource Canada's (NRCan) Precise Point Positioning (PPP) tool to post-process the GNSS observations and calculate an accurate base position using satellite orbit, clock and bias corrections. GCP locations surveyed with the rover can then be adjusted using the corrections to the base location from the PPP processing. The revised GCP locations are uploaded to the NRCan GPS-H tool to convert their ellipsoidal heights into orthometric heights, and the final coordinates exported as a csv file. The NRCan tools are free for use within the Canadian landmass, but other free online PPP options are available for use outside of Canada including; the National Geodetic Survey's Online Positioning User Service (OPUS, https://geodesy.noaa.gov/OPUS/index.jsp), GNSS Analysis and Positioning Software (GAPS) (Leandro et al., 2011), magicGNSS (Piriz et al., 2008), and the Automatic Precise Positioning Service (APPS) which uses JPL's GipsyX/RTGx software (Bertiger et al., 2020). The next pre-processing step involves refining the drone trajectory using INS data from the Applanix APX-20 UAV and converting this to orthometric heights. Working in Applanix's POSPac UAV software, the INS data are imported into a project and the base station observation added. The inertial processing tool is run to produce a smoothed best estimate of the trajectory (sbet), which is then exported in orthometric heights for import into the proprietary laser processing software (RiWorld).

We process the MiniVUX-1UAV lidar data in Riegl's proprietary RiProcess software. Each flight produces an individual laser file and associated trajectory. These are imported into the laser project and used to create an initial point cloud. The trajectories can be edited to remove unwanted data collected during turns or transits of the drone, leaving only data collected along the main survey lines. This is an important step as accelerations, decelerations, and rotations during turns and transits can negatively impact the final point cloud, creating sections with an uneven spacing of points. Once the point cloud has been sufficiently cleaned in this way, the individual flights can be merged into a single project. GCP coordinates can then be added by importing the csv file from GPS-H processing, and used to adjust and align the point clouds. In order to select the corresponding points in the laser file, we find that the cloud is best visualized using the reflectivity option. We select several points nearest the center of each GCP target and average their elevations to account for centimetric vertical scatter. Once target centerpoints have been added for several flight lines, the flight line point clouds can be precisely georeferenced, aligned, and adjusted using the RiPrecision tool. Flat hard surfaces can be used to check the amount of scatter present within the point cloud. Many of our field sites are in remote regions with limited options for additional, independent ground control, and we have therefore not performed testing of georeferencing uncertainties of our system against geodetic control monuments. However, we estimate uncertainties in the order of ∼20 cm, based on random scatter observed in our point clouds (∼15 cm) as well as expected accuracies of our GNSS ground control. We acknowledge that constraining these errors is important, especially for applications involving change detection between multiple acquisitions. However, our ∼20 cm estimate compares favourably with uncertainties associated with ALS datasets, such as errors of ∼21 cm observed by Hodgson and Bresnahan (2004) over a range of land cover types, and horizontal and vertical errors of ∼29 cm and ∼9 cm observed by Glennie et al. (2013) in helicopter lidar acquisitions.

Once the flight lines have been aligned and georeferenced, an unclassified point cloud can be exported for final classification, cleaning and gridding, for which we use a variety of programs within the licensed LAStools package (Isenburg, 2021). A copy of the shell script for classifying and rasterizing the raw lidar point cloud can be found in the supplemental material (SM12). To determine ground points we use lasground_new, a progressive morphological filter (Zhang et al., 2003). There are alternate options for ground classification, such as simple morphological filtering (Pingel et al., 2013) or cloth simulation filtering (Zhang et al., 2016), both of which can be freely used with the Point Data Abstraction Library (PDAL), an open-source library for processing and analysing point clouds (Butler et al., 2021). However, we prefer lasground_new because it has several parameters such as step, spike and bulge that can be adjusted to best find ground points according to the environment in which the data were collected, and which works well on steep, forested slopes (Căţeanu et al., 2017) that are prevalent in western Canada. Once the ground returns have been determined, we classify the remaining points using lasclassify. Other, optional steps at this stage of processing include tiling the point cloud to allow for efficient multi-threaded processing and clipping it to a polygon of interest. Isolated noise points within the cloud are then removed using lasnoise before a DTM can be generated by rasterizing the ground points with lasgrid. Lasgrid can also be used to export other desired parameters such as point density per pixel, intensity, scan angle, and (if available) RGB values. The same LAStools parameters were used to process each dataset, with the only differences being the cell size for the resultant raster products, listed in Table 3.

2.5 Data comparisons and differencing

In Sections 3–6, we compare our fully processed drone lidar data with overlapping airborne lidar and SfM surveys in order to assess their consistency and to determine from any differences whether drone lidar of-

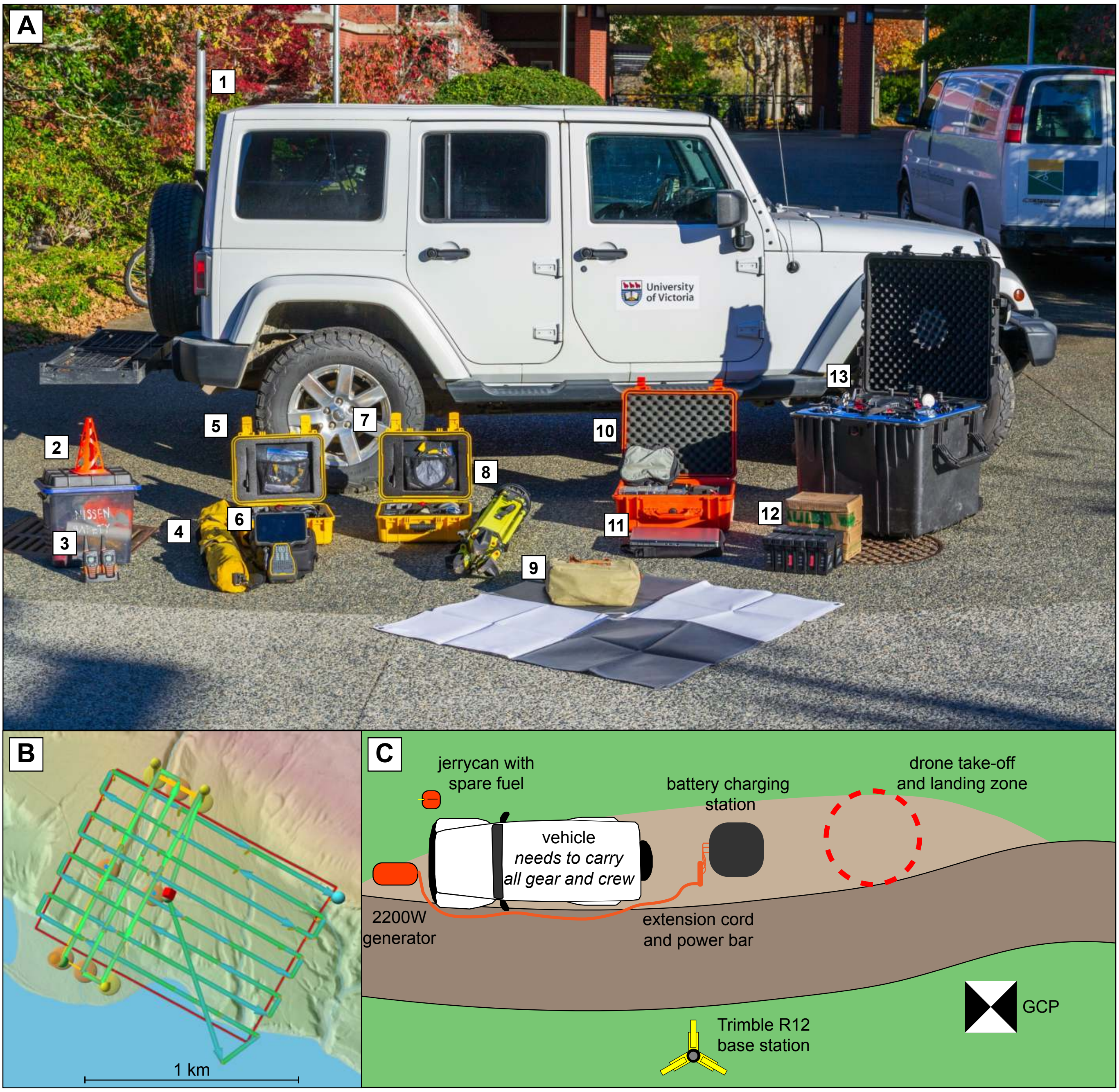


Figure 3 (A) Equipment needed for a typical acquisition. (1) Four-wheel drive vehicle. (2) Safety gear including first aid kit, traffic cones, high-visibility vests, and fire extinguisher. (3) Walkie-talkie radios for pilot and visual observers, plus chargers. (4) Trimble monopod for GNSS rover. (5) Trimble R12 GNSS rover in Pelican case. (6) TSC7 handheld computer for Trimble system. (7) Trimble R12 GNSS base station in Pelican case. (8) Trimble tripod for GNSS base station. (9) GCP targets, hammer and nails in carry bag. (10) Riegl miniVUX-1UAV laser scanner, Applanix APX-20 UAV INS, and assembly toolkit in Pelican case. (11) Windows laptop with UgCS flight planning and Riegl lidar processing software installed. (12) At least four but preferably six sets of batteries for the drone, plus two DJI Hex Chargers. (13) DJI Matrice 600 Professional hexa-copter in its customized carry-case (dimensions 68 cm $\times$ 53 cm $\times$ 49 cm). Not pictured: field iPad (to connect to the radio controller), generator (minimum 2200 W), extension cord, powerbar, and tarpaulin. (B) Flight line planning in UgCS for a segment of the Southern Rocky Mountain Trench acquisition (Section 5). The red pin is our launch site, the red box is our specified target area, and the blue-green lines are the drone flight lines. (C) Sketch-map showing a typical roadside survey set-up.

fers advantages over the more established methods. For these analyses, we first used the iterative closest point (ICP) algorithm available within free CloudCompare software (http://www.danielgm.net/cc/) to perform a final registration of the drone and comparison point clouds. The ICP algorithm finds a rigid body transformation (translation and rotation) that iteratively minimizes the closest point pair distances between two point clouds. This step was taken in order to account for several sources of error, including random scatter within each point cloud, potential differences in the global registration of the two datasets including use of different vertical datums, and/or uncertainties within the control used to georeference either dataset. Con-

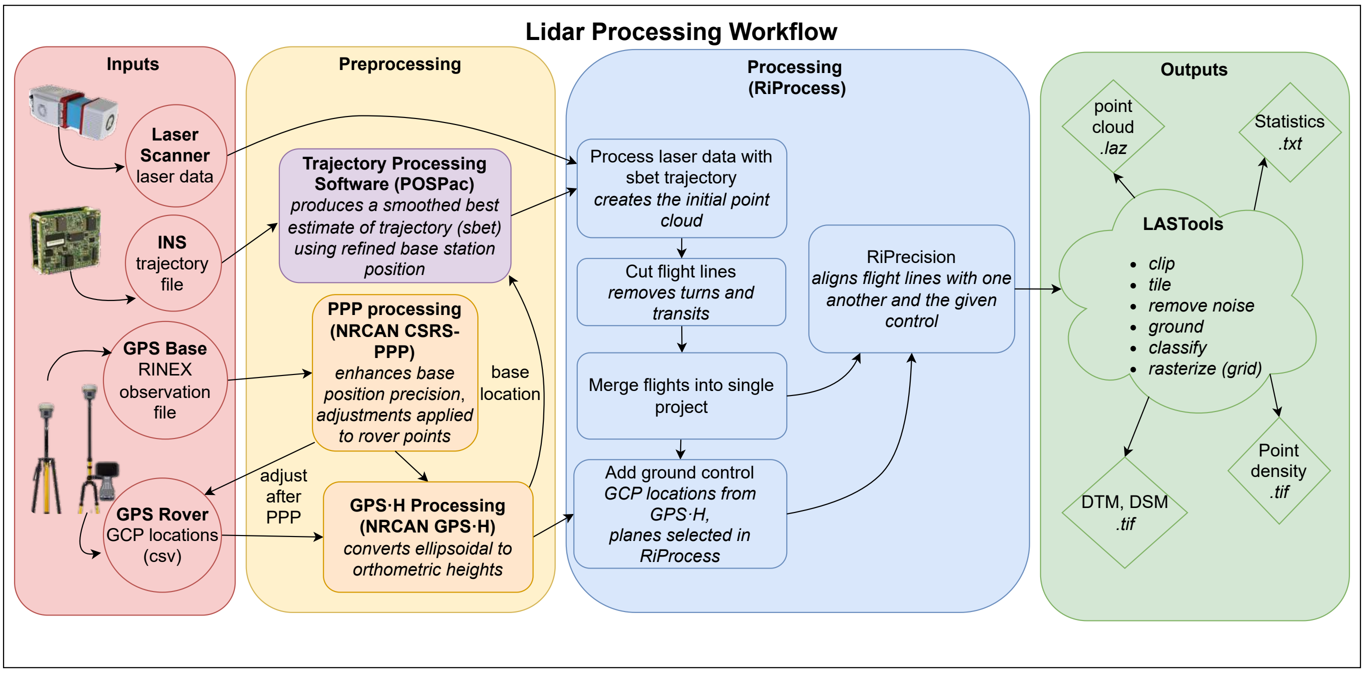


Figure 4 Flow diagram summarizing our ULS data processing workflow. Abbreviations used in this figure; DSM: Digital Surface Model, DTM: Digital Terrain Model, GPS: Global Positioning Systems, INS: Inertial Navigational System, NRCAN: Natural Resources Canada, PPP: Precise Point Positioning, sbet: smoothed best estimate of trajectory.

sequently, remaining differences between the datasets (see below) will largely represent differences in the way each method characterizes the ground surface, coupled with (likely minor or localized) natural or anthropogenic landscape change that may occurred between each pair of surveys.

We used CloudCompare's M3C2 (Multiscale Model to Model Cloud Comparison) plugin (Lague et al., 2013) to calculate the distance between ground points within the compared datasets. The M3C2 algorithm computes the local distance between two point clouds along a normal surface direction. This calculation is performed upon the point clouds, without any meshing or gridding, providing a signed 3-D distance as opposed to alternative techniques which only offer either 2-D differences (e.g. vertical difference maps) or unsigned 3-D differences (CloudCompare's cloud-to-cloud distance tool). Additionally, the M3C2 method is designed for application to datasets of contrasting point spacings (Lague et al., 2013; DiFrancesco et al., 2020), and is thus well-suited for comparing dense drone lidar datasets with coarser airborne datasets. Within CloudCompare, the drone lidar dataset was selected to be cloud #1 and the airborne lidar or SfM dataset as cloud #2. The sign of the distance reflects where the reference point cloud was along the direction of the normal for each core point. For example, a negative M3C2 value indicates an area where the drone lidar data are located underneath the airborne lidar. The cloud distances were calculated with the 'multi-scale' option, meaning that the normal distances could be in any combination of horizontal and vertical. The coarser of the two datasets was used to determine the core points as its wider point spacing reduces the number of computations required. The plugin produces a point cloud containing the M3C2 distances and areas with significant change as additional attributes. This was saved and rasterized, highlighting internal differences between the point clouds.

We further quantified disparities in the raster models produced for the drone lidar and comparison datasets by calculating their DEMs of Difference (DoD). For this, the drone lidar was first re-gridded (post point cloud alignment) at the same resolution as the airborne lidar DTM or SfM DSM, before the vertical topographic differencing was performed using the Geospatial Data Abstraction Library's gdal_warp tool (GDAL/OGR contributors, 2023). Following the methodology and conventions of Scott et al. (2021), the compare dataset (the newer drone lidar DTM) was subtracted from the reference dataset (the older airborne lidar DTM or SfM DSM). As such, negative DoD values represent areas in which the drone lidar DTM is lower than the reference ALS or SfM model, and vice versa. Note that this distance is purely vertical, unlike the M3C2 distances which are normal to the sparser of the point clouds. A shell script for differencing DTMs and generating histograms is provided in supplemental material (SM13).

3 The X̱EOLX̱ELEK̶-Elk Lake fault: a local survey of a paleoseismic trench site

3.1 Background and motivations

The X̱EOLX̱ELEK̶-Elk Lake fault (XELF) is a newly-recognized active crustal fault within the fore-arc of the northern Cascadia subduction zone on southern Vancouver Island, BC (Harrichhausen et al., 2023). The fault is named after X̱EOLX̱ELEK̶ (pronounced *hul-lakl-lik*), the name given to Elk Lake by the W̱SÁNEĆ people. The

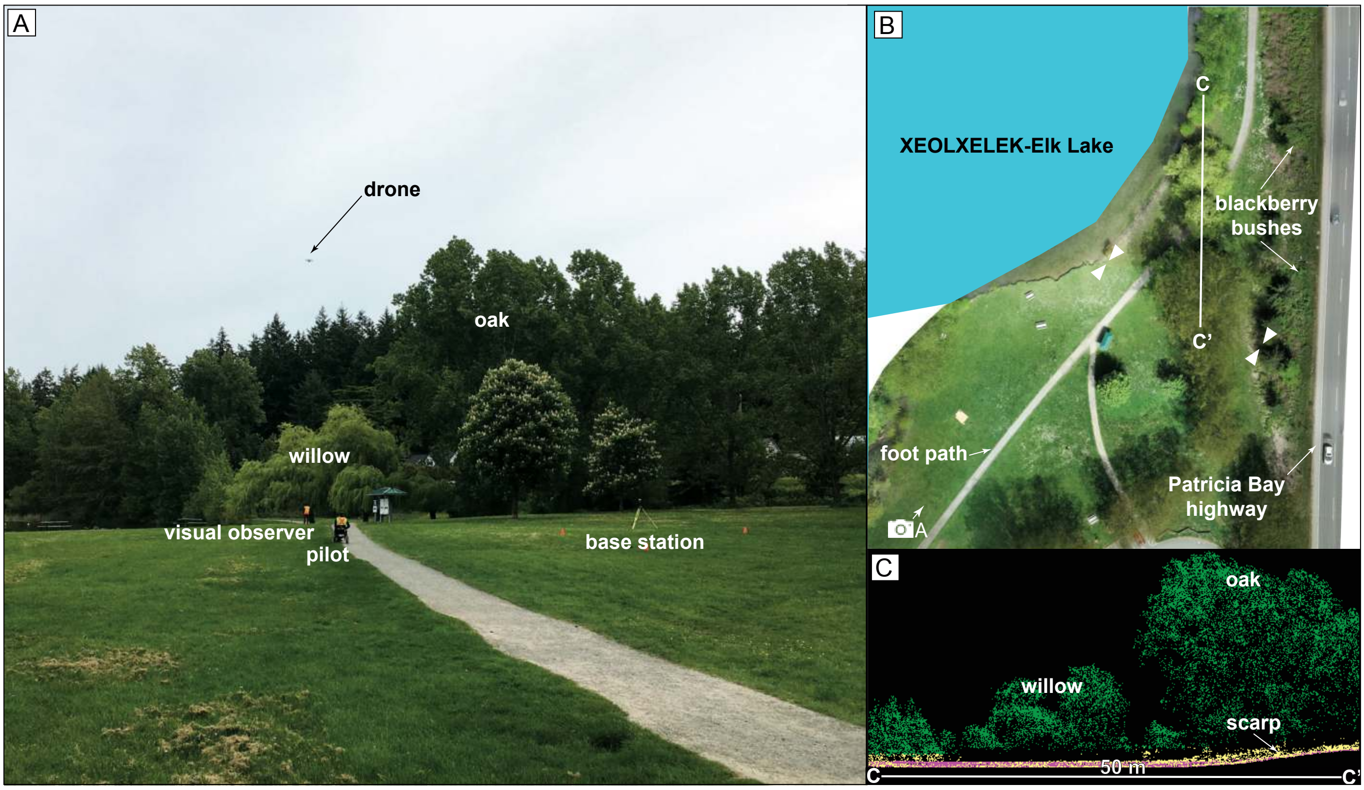


Figure 5 (A) Field photo showing ULS data collection at the eastern X̲EOLX̲ELEꝀ (Elk Lake) shoreline site. The photo location and orientation are shown in (B). The visual observer stands at the top of the XELF scarp and at the approximate mid-point of the future, NE-trending paleoseismic trench (Harrichhausen et al., 2023). (B) SfM-derived orthophoto of the acquisition site. White triangles indicate the location of the fault scarp. (C) Cross section C—C' through part of the classified drone lidar point cloud. Green points are vegetation, pink are ground points, and yellow points are unclassified.

XELF was first identified from provincial airborne lidar imagery (LidarBC, 2023) crossing Saanich peninsula between Saanich Inlet in the NW and Haro Strait in the SE (Figure 2B). The lidar data revealed several ∼N-facing fault lineaments including a 1-2.5 m high scarp displacing the surface of a Pleistocene glacial landform—a large, N–S drumlinoid ridge—between X̲EOLX̲ELEꝀ (Elk Lake) and Haro Strait. A site was chosen for paleoseismic trenching between the eastern lake shoreline and the Patricia Bay Highway, where the scarp passes through Elk-Beaver Lake Regional Park. The trench, excavated in August 2021, contained evidence that indicated the XELF has ruptured in at least one large (M_w ∼6.1–7.6) thrust earthquake during the late Holocene (Harrichhausen et al., 2023).

The eastern X̲EOLX̲ELEꝀ (Elk Lake) shoreline site (Fig. 5A–B) provided an early and relatively simple test of our new ULS system, acquired prior to the planned trench and providing the best possible data for measuring the local scarp height. The survey area has dimensions of just ∼100 m and exhibits gentle relief other than a steep bank up to the Patricia Bay Highway along the eastern boundary. The ground cover is mostly mowed grass as well as scattered blackberry bushes, deciduous trees (willow and oak) and conifers (Douglas Fir and Western Red Cedar), the tallest of which are around 35 m. Since the site lies within the municipal Elk-Beaver Lake Regional Park and on traditional W̱SÁNEĆ territory, further permissions had to be obtained to conduct research and operate a drone within the park, in addition to the civil aviation approvals described in Section 2.2 (see Acknowledgments).

3.2 ULS data acquisition and results

We surveyed the eastern X̲EOLX̲ELEꝀ (Elk Lake) shoreline site with our drone lidar system in May 2021 (Fig. 5A), three months prior to the paleoseismic trench excavation by Harrichhausen et al. (2023). Our 100 m × 110 m (∼11,000 m^2) drone lidar dataset took a total of 4 hours to collect, including set-up, with a crew of a pilot, two visual observers, and two assistants who helped avoid flying over pedestrians in the busy park (see Section 2.1). The drone was flown at a height of 45 m AGL in a cross-hatch pattern of N-S and E-W flight lines (supplementary Figure SM1) and at a relatively slow speed of 2 m s^{-1}. These parameters were chosen in order to collect as high a resolution dataset as possible, with an expectation of >100 pts/m^2. At 45 m AGL the laser footprint is 3.6 cm at the center of each 90 m-wide swath and 5 cm at the edges of the swath. Three GCPs were deployed and used to georeference the dataset.

Processing and classifying the drone lidar data using the workflow described in Section 2.4 yielded an average point density of 543 pts/m^2 and an average classified ground return density of 260 pts/m^2, leading to an average ground return spacing of ∼6 cm (Figure 6A–B, middle column, and Table 3). Classified ground returns constitute ∼48% of all points in the cloud. As expected, ground returns are densest away from the

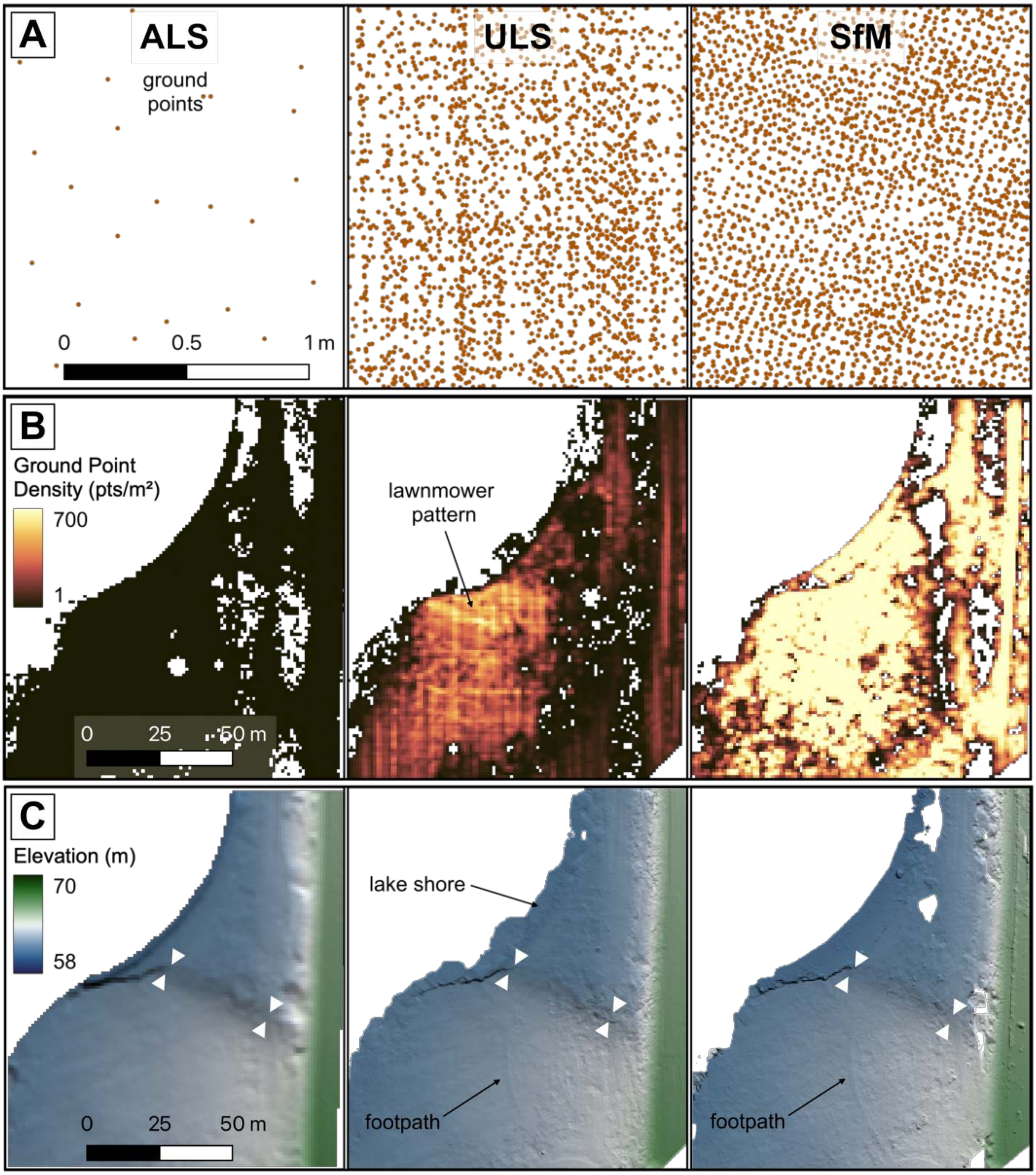


Figure 6 Comparison of topographic datasets at the eastern X̱EOLX̱ELEꝴ (Elk Lake) shoreline site: (left column) LidarBC ALS data, (middle) our ULS data, and (right) our SfM data. (A) Representative ground point clouds viewed at a roughly metric scale in order to contrast point spacings. (B) Survey-wide classified ground point densities. The same color palette is used for each plot, and white spaces indicate areas without ground coverage (mostly trees and dense blackberry bushes). (C) Hillshaded DTMs illuminated from the SSW (210°) in order to highlight the NNE-facing fault scarp, further delineated by white arrows. The DTMs were constructed with an interpolation of 5 pixels in order to minimize holes.

trees over areas of open, mowed grass, reaching values as high as $\sim$700 pts/m^2 where swaths from several flight paths overlap. However, a visual inspection of the classified point cloud in cross section also reveals successful imaging of the ground surface through tree and shrub foliage (Fig. 5C). We find that a cell size of 20 cm optimizes the ULS raster DTM, minimizing its pixel dimensions without introducing widespread data gaps (Fig. 6C, middle column). The cross-hatched point density pattern (Fig. 6B, middle column) results from

General Acquisition Information

Paper section	Fault (site)	Data type	Provider	Area (km^2)	Point density (pts/m^2)	Ground return density (pts/m^2)	Ground return spacing (m)	% pts ground	DTM cell size (m)
3.2	XELF	ULS	This study	0.01	543.39	260.11	0.06	48	0.2
	XELF	ALS	LidarBC		13.97	9.00	0.33	64	1.0
	XELF	SfM	This study		1906.16	554.04	0.04	29	0.2
4.2	SJF	ULS	This study	2.81	130.28	13.38	0.29	10	
	east			1.21	141.70	10.80	0.30	7	0.5
	west			1.59	121.59	17.24	0.24	13	0.5
	SJF	ALS	Mosaic		(ground only)	5.70	0.43	-	1.0
5.2	SRMT	ULS	This study	3.14	101.57	35.62	0.17	35	0.3
	SRMT	ALS	LidarBC		17.71	7.00	0.38	41	1.0
6.2	EDF	ULS	This study	10.42	97.54	45.47	0.21	47	
	BURW			0.37	92.75	25.95	0.20	28	0.3
	COPJ			1.02	79.60	25.46	0.20	32	0.3
	DUKE			4.86	112.33	71.78	0.19	64	0.3
	NINE			0.82	90.33	21.43	0.22	24	0.3
	QUIL			1.34	96.96	21.45	0.22	22	0.3
	SLIM			0.52	82.25	29.00	0.29	35	0.3
	TELL			1.49	81.61	19.69	0.23	24	0.3
	EDF	ALS	This study		7.85	3.45	0.54	44	1.0

Paper section	Fault (site)	Data comparison	ICP transformation: Max rotation (°)	ICP transformation: Translation vector (m)	Differencing M3C2: Mean (m)	Differencing M3C2: SD (m)	Differencing DoD: Mean (m)	Differencing DoD: SD (m)
3.3	XELF	ULS–ALS	0.0008	0.28	−0.05	0.19	−0.32	0.26
	XELF	ULS–SfM	0.0005	0.33	−0.03	0.14	−0.27	0.27
4.3	SJF (east)	ULS–ALS	0.0001	0.07	0.22	0.42	0.21	0.62
	SJF (west)	ULS–ALS	0.0002	0.32	0.13	0.31	0.15	0.53
5.3	SRMT	ULS–ALS	0.0001	0.54	−0.01	0.09	−0.02	0.29
6.3	EDF (BURW)	ULS–ALS	0.0004	0.82	−0.01	0.14	0.01	0.26
	EDF (COPJ)	ULS–ALS	0.0002	0.87	−0.01	0.17	0.00	0.25
	EDF (DUKE)	ULS–ALS	0.0003	0.54	0.00	0.18	0.00	0.28
	EDF (NINE)	ULS–ALS	0.0002	0.99	0.01	0.18	0.01	0.23

Table 3 Statistics of our ULS and comparison datasets (upper part of the table) and of the ICP alignments and subsequent M3C2 and DoD differencing (lower part of the table). ALS statistics are calculated from within the footprints of the corresponding ULS surveys, allowing a like-for-like comparison. % pts ground is the percentage of all lidar returns classified as ground.

the perpendicular orientations of the survey flight lines.

3.3 Comparisons and differencing with ALS and SfM data

Our first comparison dataset is the provincial airborne lidar survey, flown in 2019 and available from the LidarBC (2023) portal. Within the ULS survey footprint, the ALS point cloud yields an average point density of 14 pts/m^2, an average ground return density of 9 pts/m^2 for a spacing of ∼33 cm, and a maximum ground return density of 10 pts/m^2 (Figure 6A–B, left column, and Table 3). The ALS point cloud also contains more, and larger, data gaps than the ULS cloud, indicating greater difficulty in imaging beneath vegetation. The ALS DTM has a cell resolution of 1 m, 25 times coarser than our ULS DTM (Figure 6C, left column). The fault scarp can be made out in both raster datasets as a gentle NNE-facing slope trending WNW-ESE across the center of the target area, but the finer resolvability of the ULS dataset

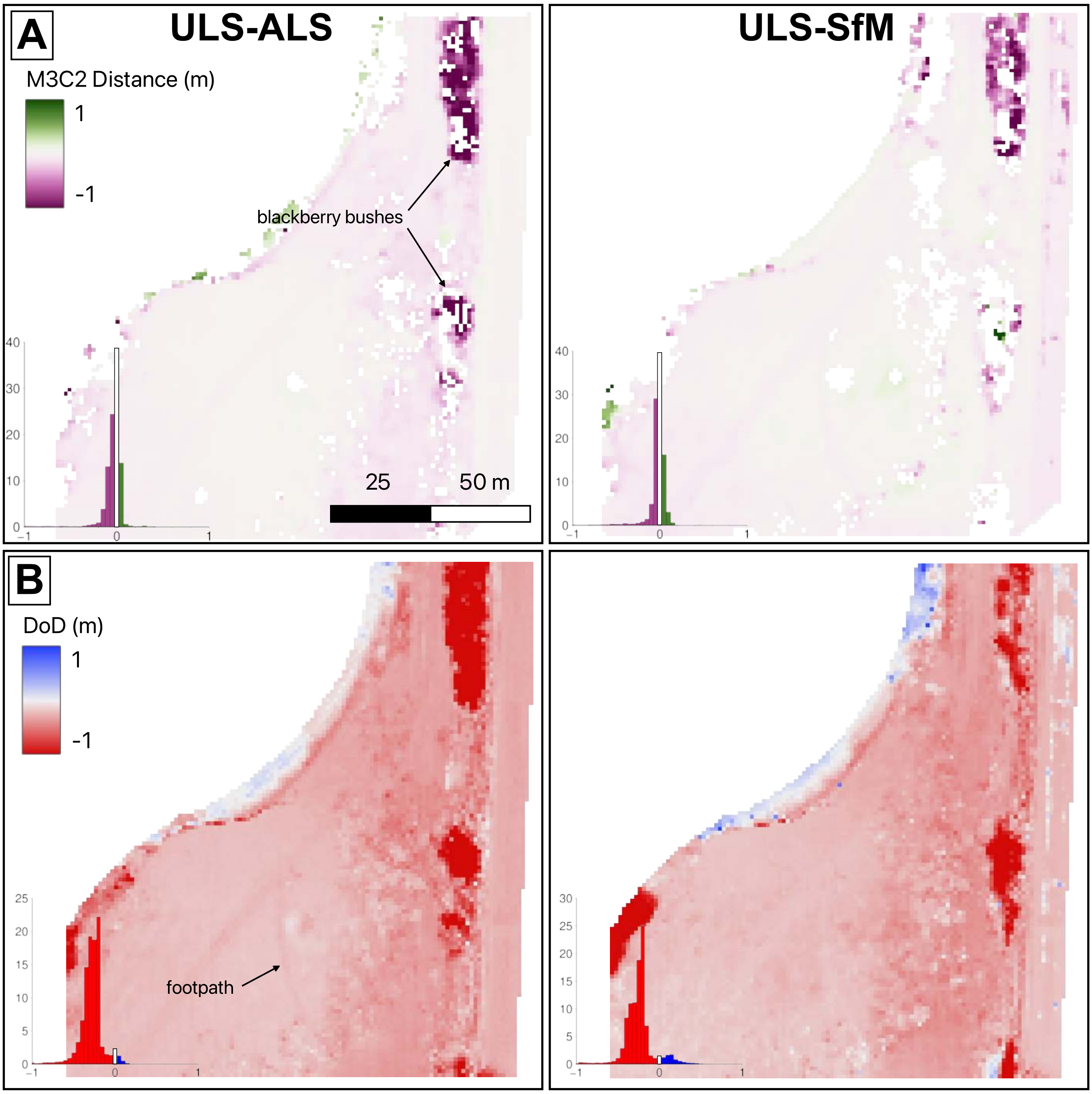


Figure 7 Differencing of the ULS and ALS datasets (left hand column) and of the ULS and SfM datasets (right hand column) at the eastern X̱EOLX̱ELEK̶ (Elk Lake) shoreline site, shown as (A) M3C2 distances calculated in CloudCompare and (B) a DEM of Difference. Positive values indicate where the ULS dataset was higher than the comparison dataset. Histograms show distributions of raster values.

is evident in a small linear depression along one of the park footpaths, which is not visible in the ALS DTM. Though not a tectonic feature, this does highlight the potential for ULS to identify subtler fault offsets than are evident in traditional airborne lidar data.

Because much of the eastern X̱EOLX̱ELEK̶ (Elk Lake) shoreline site is covered by mowed grass, it provides our best opportunity out of all of our case studies to compare our drone lidar with SfM data. With this in mind, we surveyed the site with SfM two weeks before our drone lidar flights. Using a DJI Phantom 4 Professional V2 drone with the built-in camera, we captured 749 photographs which were then processed using the Agisoft Metashape Professional software package, with 15 GCPs deployed across the scene for georeferencing (e.g. Johnson et al., 2014). After classifying the SfM point cloud with the LAStools LASclassify program, we yielded an average point density of 1,906 pts/m^2 and average classified ground point density of 554 pts/m^2 and spacing of 0.04 m (Figure 6A–B, right column, and Table 3). However, the SfM is strikingly less uniform than either lidar point clouds, containing far more, and larger, data gaps, reflecting its inability to image beneath vegetation (Figure 6B). As a consequence, we find that the optimal resolution of the SfM DTM is 20 cm (Figure 6C, right column), no finer than the ULS DTM despite their differing underlying ground point densities.

The ICP rigid body transformation that most closely aligns the ULS point cloud to the ALS point cloud involved a translation vector of 28 cm and rotations of <0.001 radians (Table 3). These values reflect small differences, within the expected error, of the global registration of the two surveys, with minimal tilting of one relative to the other. The mean M3C2 distance between the airborne and drone lidar point clouds was 5 cm with a standard deviation of 19 cm (Fig. 7A, left panel). The equivalent DEM of Difference (DoD) exhibits a mean of −0.32 m and a standard deviation of 26 cm (Figure 7B, left panel). Positive values (blue colours) reflect areas where the ULS DTM is higher than the ALS DTM, and negative values (red colours) reflect those where the ULS DTM is lower. These non-zero values reflect a combination of factors, including residual misalignment of the point clouds (even after ICP co-registration) and internal vertical scatter, estimated from hard, flat, non-vegetated surfaces (e.g. roads, parking lots) at ±6.7 cm in the ALS point cloud and ±7.5 cm in our ULS point cloud. However, careful analysis of the DoD and M3C2 maps also supports a third cause of vertical differences. The largest differences (−2.6 m) were negative and are most likely the result some areas where returns off of dense vegetation were misclassified as ground in either of the datasets. These highest values are approximately in the same area that is covered by dense blackberry bushes (Fig. 5). This also coincides with areas in the airborne lidar that have few to no points (Fig. 6C). It is likely that the airborne dataset has its lowermost returns from within the blackberry bush rather than the ground surface, while the drone lidar, with its denser point cloud has managed to capture a better ground surface beneath these dense bushes. This also explains why the M3C2 distances are largest in the very same areas.

The ICP rigid body transformation that aligns the ULS and SfM point clouds involved rotations of <0.01 radians and a translation vector of 33 cm, again indicating consistency to within a few decimeters in global registration of the two datasets. The mean M3C2 distance was −0.03 m with a standard deviation of 27 cm (Fig. 7A, right column), while the DoD had a higher mean value of −0.27 m and a standard deviation of 27 cm (Fig. 7A, right column). Given the stark contrast in vegetation penetration capability, it is difficult to interpret these centimetric-to-decimetric differences between the ULS and SfM surveys. Similar to the airborne lidar comparison, the greatest discrepancies are in areas that are covered in dense bushes, where the ULS alone seems to penetrate to ground level.

4 The San Juan fault: a kilometric survey of a fault scarp in steep, forested terrain

4.1 Background and motivations

The San Juan fault (SJF) is a major crustal fault on southern Vancouver Island, located north-west of the XELF (Fig. 2B). The SJF transects the island west to east for ∼80 km across densely forested hills of the southern Vancouver Island ranges. In our area of interest, the fault separates the intrusive West Coast Crystalline Complex from the extrusive Jurassic Bonanza Group of the Wrangellia terrane (Harrichhausen et al., 2022). There have been numerous interpretations, inferred from regional geology, of the roles that the SJF has played throughout its evolution (Johnson, 1984; Rusmore and Cowan, 1985; Brandon, 1989; England and Calon, 1991). Most recently, its kinematics have been constrained to have been left-lateral during Eocene accretion of the Crescent-Siletz terrane (Harrichhausen et al., 2022). The position of the SJF in the forearc of the active Cascadia subduction zone, its favourable orientation relative to the regional stress field, and its conspicuously linear trace motivate a close examination of its current activity, but no convincing evidence of recent earthquake rupture has yet been found. If the SJF is active, it may pose a considerable risk to Victoria, Nanaimo, and other towns and infrastructure along Vancouver Island's Highway 1 corridor.

For our drone lidar surveying of the SJF, we targeted a ∼4 km section of the fault accessed via logging roads west of Shawnigan Lake (Fig. 8A). The SJF trace is locally defined by a N-facing scarp that appears to cross-cut a glacially scoured surface as well as a number of small tributaries of the Koksilah River. The presence (or absence) of faulted offsets to these glacial and fluvial features could help determine whether this section of the SJF has been active in the late Quaternary. The area of interest includes some steep slopes with a topographic variation of 300 m. Vegetation cover includes stands of second growth Pacific cool temperate forest (Baldwin et al., 2019), with Douglas firs, Western Red Cedars, Western Hemlock and Sitka Spruce trees that are up to 50 m tall. Additionally, some of the area includes recent clearcuts with some small trees and shrubs. Our surveying of the SJF therefore provides good tests both of mapping at kilometric lengthscales over rugged terrain and of the vegetation penetration capability of the drone lidar system.

4.2 ULS data acquisition and results

We surveyed the SJF with our lidar drone in September 2022. For logistical reasons, we split the target area into two sections separated along strike of the SJF by a gap of ∼1 km (Fig. 8A). The SJF West section is ∼2 km along strike by ∼700 m wide and the SJF East section is ∼1.7 km along strike by ∼700 m wide, with a total surveyed area of 2.9 km^2. The two sites took a total of 2 days to survey (14 hours of field work excluding travel) with a crew of three people (a pilot and two visual observers). The drone was flown at 5 m s^{-1} at a height of 80 m AGL along flight lines oriented parallel to the mapped fault scarp, that were merged with the help of orthogonal calibration lines (supplementary Figures SM2 and SM3). For the SJF-East section, two launch points were required to maintain VLOS around high topography in the center of the survey area. At 80 m AGL the laser footprint is 6.4 cm at the center and 9 cm at the edges of each ∼160 m wide swath swath. At each site we deployed 5

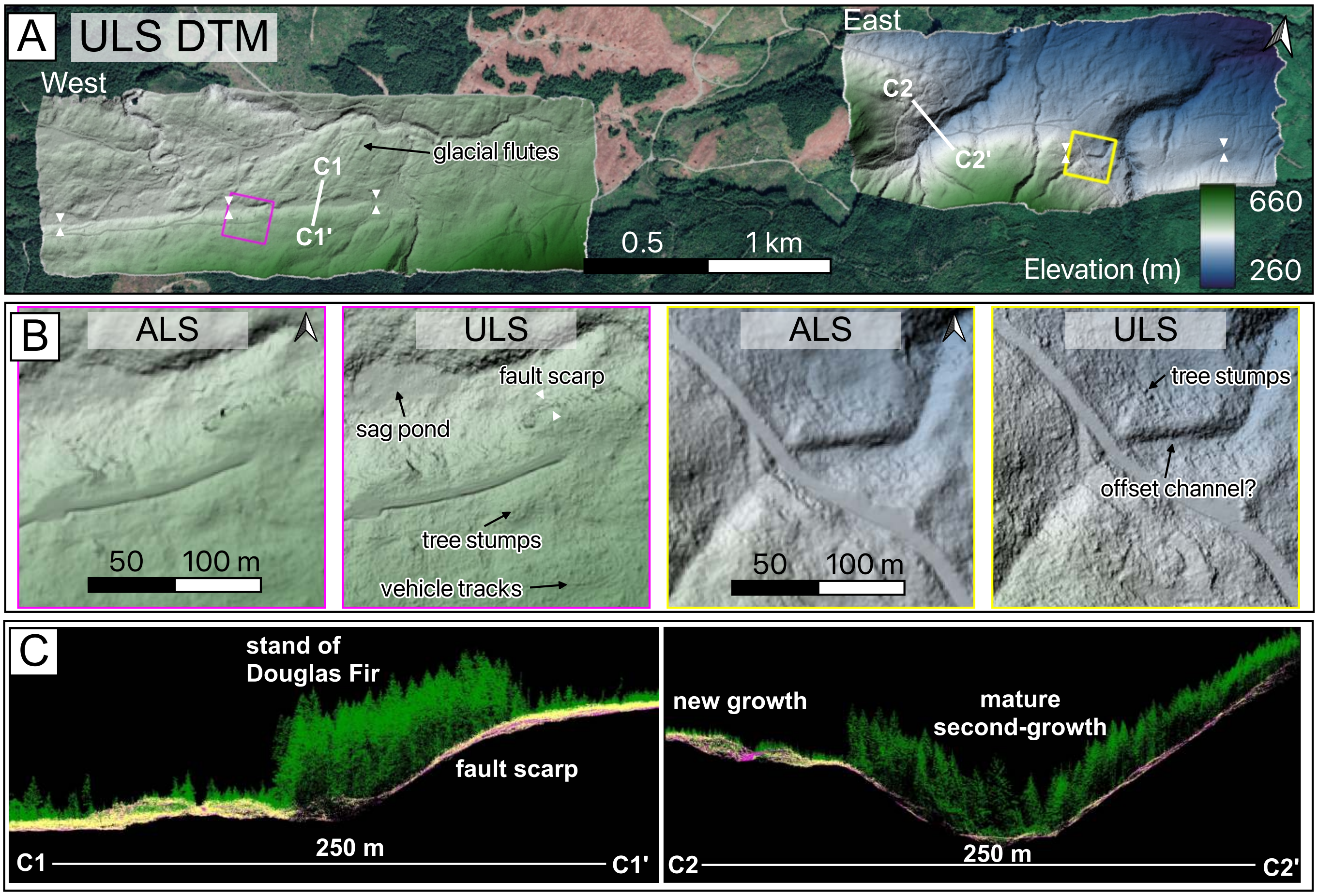


Figure 8 (A) Hillshaded UAV lidar DTM (illuminated from 315°) for the San Juan Fault study area, overlain on satellite photograph. White triangles indicate the approximate location of the SJF. Coloured squares indicate the areas shown in the DTM comparisons below. (B) Comparisons between ALS and ULS DTM hillshades at the SJF West site (left hand panels) and the SJF East site (right hand panels). (C) Cross sections C1–C1' and C2–C2' through the ULS classified point cloud. Green points are vegetation, pink are ground points, and yellow points are unclassified.

GCPs to assist with georeferencing the point clouds.

Our SJF East and SJF West ULS surveys have average point densities of 121 pts/m^2 and 142 pts/m^2 and average classified ground return densities of 11 pts/m^2 (0.3 m spacing) and 17 pts/m^2 (0.24 m spacing), respectively (Table 3). The point clouds are therefore an order of magnitude sparser than those at X̲EOLX̲ELEK̶ (Elk Lake), reflecting the greater platform heights and speeds and reduced swath overlap used in our deployments along the SJF. Additionally, only ~10% of the laser returns along the SJF are classified as ground compared to ~48% at X̲EOLX̲ELEK̶ (Elk Lake), reflecting the stark differences in vegetation between the two target areas. Nevertheless, the drone lidar still captures an abundance of ground surface returns along the SJF, from beneath both mature forest and new growth within clear cuts (Fig. 8C). The optimized 0.5 m-pixel hillshaded ULS DTM captures clearly both the ~E–W-striking fault scarp and several ~NNE–SSW-trending glacial flutes at the SJF West site, as well as the ~NE-trending tributary channels at the SJF East site, one of which exhibits an apparent right-lateral offset at the fault (Fig. 8A). Further analysis and interpretation of this rich dataset will form the basis of future study.

4.3 Comparison and differencing with ALS data

We can compare our ULS data with regional airborne lidar flown for Mosaic Forestry Management in 2021. In the year between the two surveys there appears to have been little forestry activity in the area (and no new cut blocks), allowing a like-for-like comparison. Within the footprint of the ULS surveys, the ALS yields an average ground point density of 6 pts/m^2 and average spacing of 0.43 m, somewhat coarser than the ULS data. A visual comparison of the 0.5 m-pixel ULS DTM with the 1 m-pixel ALS DTM demonstrates how the drone lidar allows for finer scale (<1 m) features to be identified (Fig. 8B). For example, tree stumps and vehicle tracks on clear-cut slopes are clearly visible on the drone lidar hillshade but are only vaguely delineated in the airborne lidar.

The ICP rigid body transformations that best aligned the ALS and ULS point clouds involved rotations of <0.001 radians and translation vectors of 0.07–0.32 m. Post alignment, the average M3C2 distances were 0.22 m for SJF East and 0.13 m for SJF West, with standard deviations of 0.42 m and 0.31 m, respectively, while the equivalent DoDs have mean elevation discrepancies of 0.21 m and 0.15 m with standard deviations of 0.62 m and 0.53 m, respectively (Fig. 9 and Table 3). These re-

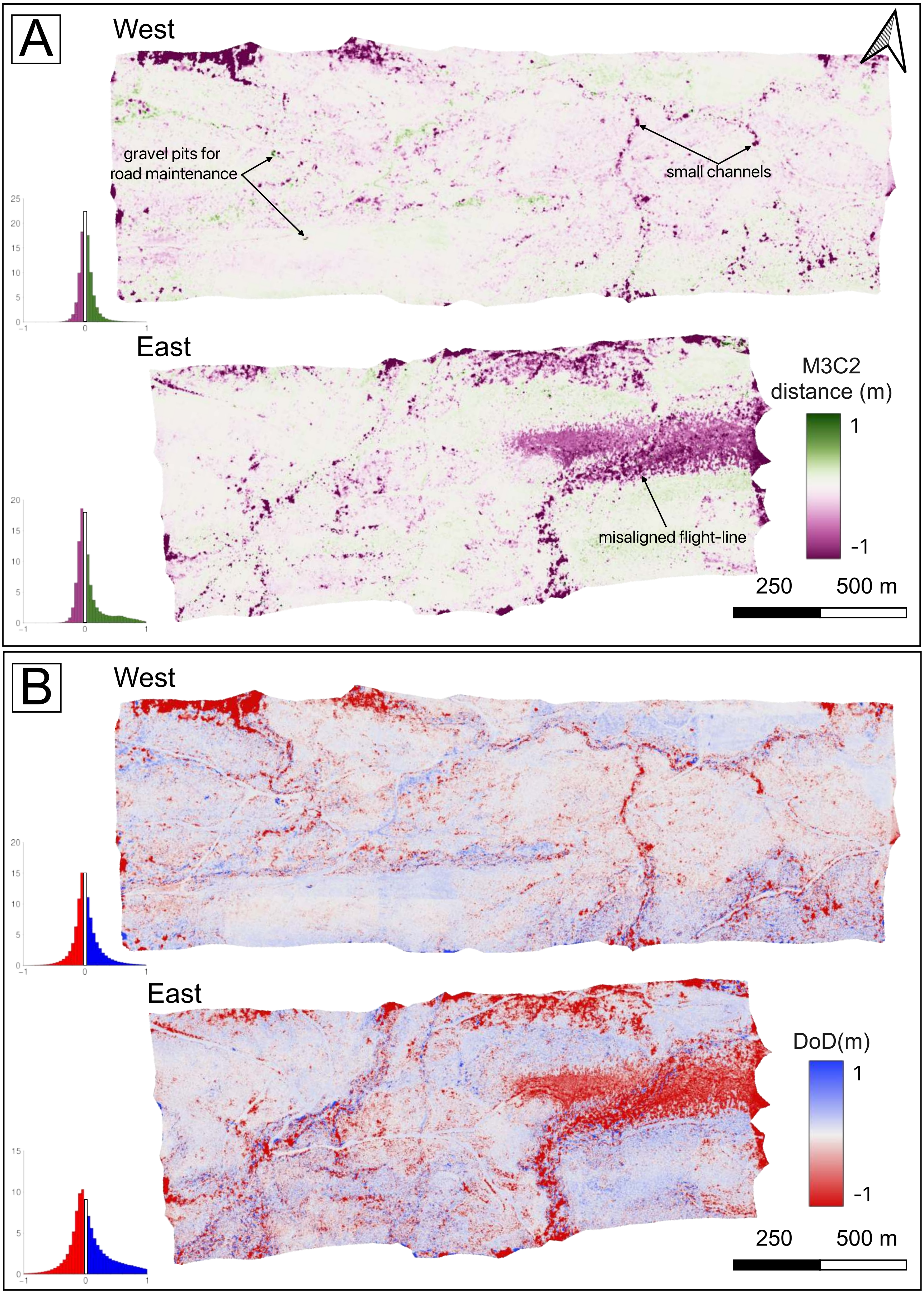


Figure 9 Differencing of the drone and airborne lidar datasets for the SJF study area, shown as (A) M3C2 distances calculated in CloudCompare and (B) a DEM of Difference. Positive values indicate where the ULS dataset was higher than the ALS dataset. Histograms show distributions of raster values.

sults indicate internal consistency of the two datasets to within a few decimeters.

The largest M3C2 distances in the SJF West dataset occur along the northern edge of the ULS survey within a steep valley (Fig. 9A). Our point cloud is sparsest in this area, as it was covered fully by just one flight line. Other small areas with large M3C2 distances highlight where excavations for road maintenance were made between acquisitions (Fig. 9A). There are also small channels that show up as negative values in the M3C2 distance and DoD plots, as a result of improved penetration through dense riparian vegetation in the ULS dataset. The strip of high M3C2 distances in the eastern part of the SJF East dataset results from a mis-aligned ULS flight line, which we discuss further in Section 7.2. In general, the raster differences are a lot noisier than the M3C2 point cloud comparison (Fig. 9B). The largest raster differences are concentrated at the bottom of valleys, areas with both steep slopes and dense vegetation. It is likely that fewer true ground returns were obtained in these areas, but particularly in the ALS dataset, as the ULS dataset generally places the valley floors lower. Similar to the data comparisons undertaken in the previous section, this highlights the better vegetation penetration capability of the drone lidar system.

5 The Southern Rocky Mountain Trench: a kilometric survey of an alluvial fan scarp

5.1 Background and motivations

The Rocky Mountain Trench (RMT) is a conspicuously linear series of NW-trending valleys that crosses the Canadian Cordillera from northern Montana to southern Yukon, where it continues as the Tintina Trench into Alaska (Fig. 2A). It demarcates the boundary between the Omineca and Foreland morphogeological belts and is defined by a series of major fault zones with distinct northern, central and southern segments (Clague, 1975; Gabrielse et al., 1991). The Southern RMT fault (SRMTF), in the East Kootenay region of southeastern BC is a steeply west-dipping normal fault active primarily in the Eocene (van der Velden and Cook, 1996). However, there is some evidence that the SRMTF may remain seismically active (Purba et al., 2021; Finley et al., 2022b), strongly motivating the acquisition and interpretation of lidar data. Our preliminary analysis of newly-released provincial airborne lidar (LidarBC, 2023) revealed a ∼3 km-long, W-facing scarp crossing a series of potentially Holocene-aged alluvial fans above the eastern shoreline of Columbia Lake, just south of the town of Fairmont Hot Springs (Fig. 10A). The exact trace of the SRMTF is not well mapped at this location owing to the thick overburden in the valley floor. However, the scarp is parallel to and aligned with mapped strands of the SRMTF to the north and south, and could potentially indicate a neotectonic reactivation. Given that this part of the RMT was occupied by Glacial Lake Invermere during the late Pleistocene (Sawicki and Smith, 1992), other potential origins including wave-cut shorelines or slumping within weak glaciolacustrine sediments must also be considered.

Our goal in surveying the Columbia Lake scarp with our lidar drone was to help determine its true origin. This includes illuminating any lateral offsets to a series of small runnels that cross the scarp, and characterizing its detailed shape for the purpose of morphologic dating (e.g. Nash, 1980; Arrowsmith et al., 1998; Hilley et al., 2010) or to reveal any bevels that might indicate a compound, multi-earthquake origin (e.g. Zhang et al., 1986; Johnson et al., 2018; Wei et al., 2019). The alluvial fan that constitutes our principle target forms a gently-sloping surface from the western front of the Stanford Range at ∼900–1,000 m elevation to the lake shoreline at ∼809 m. This is covered by a mix of open grassland and groves of ponderosa pine, typical of Cordilleran dry forest (Baldwin et al., 2019). The Columbia Lake site therefore provides a test of our ULS system across a gentler relief and more sparsely-vegetated landscape than along the SJF. Since the survey area lies within traditional territories of the Ktunaxa and Secwépemc First Nations as well as within the Columbia Lake Provincial Park and Nature Conservancy of Canada land, extra research and drone use permissions had to be obtained in addition to the civil aviation approvals described in Section 2.2 (see Acknowledgments).

5.2 ULS data acquisition and results

We surveyed the Columbia Lake site over a period of three days in October 2022 with a crew of two people (the pilot and one visual observer). Our survey covers ∼3.22 km^2, with a length of ∼4 km along strike of the scarp and a width of ∼0.8 km, enough to capture most of the fan surfaces between the mountain rangefront and the Columbia Lake shoreline. The target area was flown in several segments, with orthogonal calibration lines that tie the flights together (supplementary figure SM4). Launch sites were located along the park access road that conveniently runs N–S down the middle of the fan surfaces, often adjacent to the scarp itself. The drone was flown at a height of 80 m AGL and a speed of 4 m s^{-1}, with the gentle relief and mix of grassland and scattered ponderosa pine allowing for excellent sight lines. The ULS data were georeferenced using ten harlequin-iron-cross GCPs.

Our ULS system yielded an average point density of 102 pts/m^2 and was easily able to image beneath the scattered ponderosa pine trees (inset, Fig. 10), producing an average classified ground return density of 36 pts/m^2 at an average spacing of 0.17 m (Table 3). Overall, ∼35% of all laser returns are classified as ground, lower than the ∼48% at X̱EOLX̱ELEꝄ (Elk Lake) but substantially higher than the ∼9% along the SJF, reflecting the differing vegetation densities of the three areas. We optimally gridded the classified ground returns at a pixel resolution of 30 cm (Fig. 10A). The hillshaded DTM clearly reveals the primary scarp striking N–S across the largest, northern alluvial fan (af1), as well as some secondary splays just east of it (Fig. 10D, right panel). It also reveals a lineament within the southernmost alluvial fan (af4) that may represent an along strike continuation of the scarp. We encountered

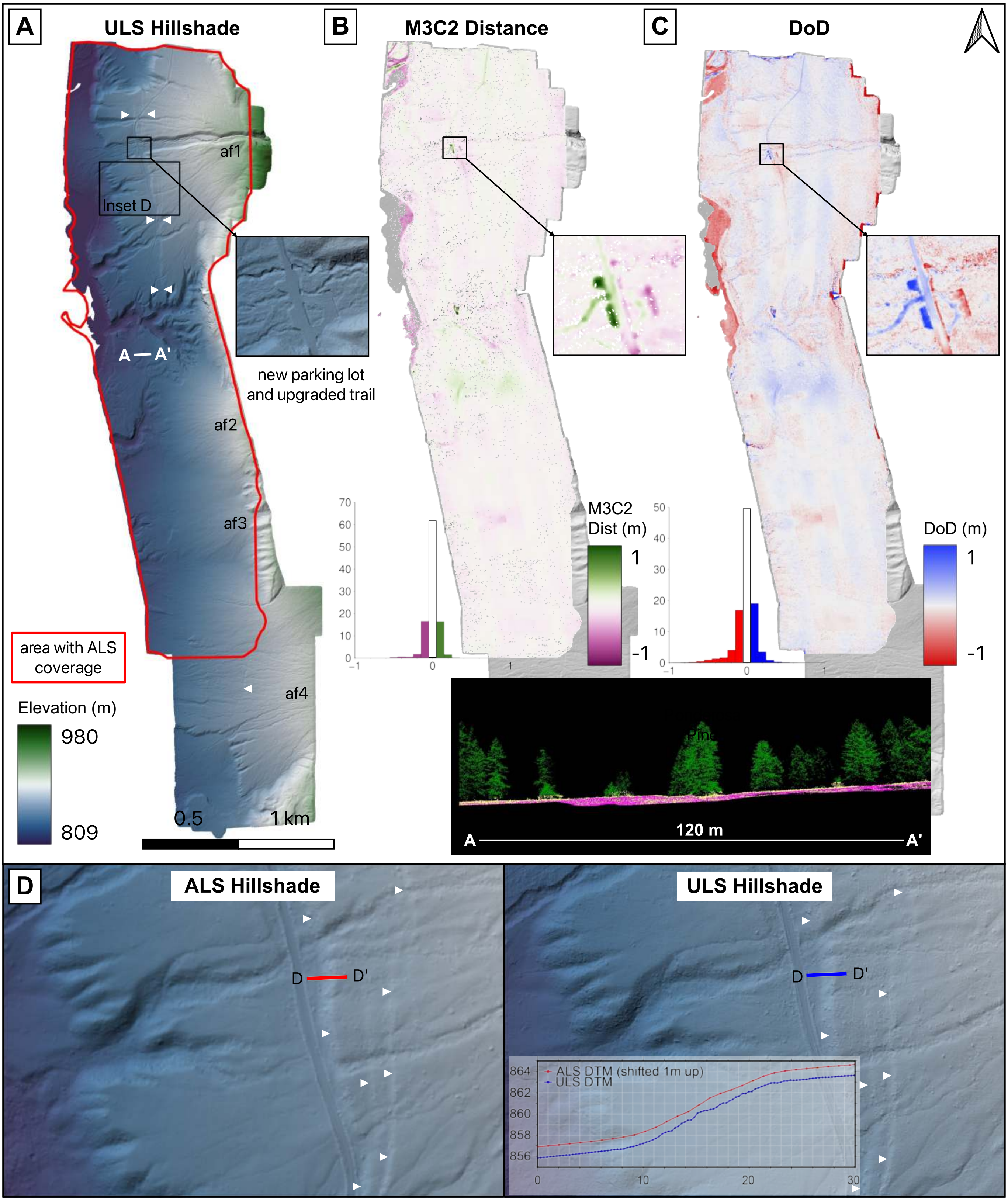


Figure 10 (A) Hillshaded ULS DTM for the Columbia Lake site along the SRMTF. The red polygon shows the area of the ULS survey overlapped by LidarBC ALS coverage. Inset shows the location of a new parking lot and upgraded trail to the lake shore that was not developed when the provincial lidar was collected. Alluvial fans (af) are numbered from northernmost (af1) to southernmost (af4). White triangles indicate the approximate location of the SRMTF. (B) M3C2 distances between the ULS and ALS point clouds, with an inset showing changes over the new parking lot and trail. Positive values indicate where the ULS dataset was higher than the ALS dataset. (C) DoD (ULS-ALS), with an inset showing changes over the new parking lot and trail. The inset below (C) and (D) shows cross section A—A' through the classified drone lidar point cloud, with green points for vegetation, pink for ground points, and yellow for unclassified returns. (D) Comparison between ALS and ULS hillshaded DTMs. Cross section D-D' shows the increased level-of-detail in the ULS DTM along the main scarp. Note that the ALS profile has been shifted upwards by 1 m in order to aid comparison.

difficulties aligning some of the flight lines due to poor INS calibration, which may explain some N-S linear corduroy artefacts visible in the center of the hillshaded DTM (Fig. 10D, right panel). However, to the trained eye, these minor and localized artefacts are easily distinguished from genuine tectonic landforms.

5.3 Comparison and differencing with ALS data

Our comparison data are provincial airborne lidar collected over a two year period (2015–2017) using multiple sensor platforms with unknown acquisition parameters (LidarBC, 2023). The sparse metadata owe to the fact that these surveys were flown by a third party and later acquired for LidarBC, without the control needed for them to verify accuracies. This captures $\sim$ 80% of our ULS dataset, but misses the southernmost alluvial fan (af4) surveyed with the drone, where we observe an additional scarp along strike. Within the footprint of overlap, the ALS data have an average ground return density of 7 pts/m^2, five times coarser than the ULS survey, and an average ground return spacing of 0.38 m, twice that of the ULS survey (Table 3). Despite these differences in spatial resolution, there is little visual contrast between the ALS and ULS hillshaded DTMs (Fig. 10D). However, fault-perpendicular topographic profiles reveal that the shape of the scarp is captured by ULS at greater detail than by ALS (inset to Fig. 10D).

The ICP rigid body transformation that most closely aligns the two point clouds has a translation vector of 0.54 m and a rotation of <0.001° (Table 3). After this global registration, the mean M3C2 distance between the aligned point clouds is -0.01 m with a standard deviation of 0.09 m, while the DoD has an average elevation difference of -0.02 m with a standard deviation of 0.29 m. These results show that despite some artefacts (described below), our ULS dataset is still both aligned to within a few decimeters and internally consistent to within a few centimeters with the ALS survey.

The greatest differences between the two point clouds as revealed by M3C2 and DoD maps (Fig. 10 B–C) are along the park access road: a gravel pit in the south and a parking lot in the north, both of which we suspect were excavated or re-graded between surveys. There are also negative M3C2 and DoD values along the Columbia Lake shoreline, which likely represent changing water levels between acquisitions. However, the M3C2 distance and DoD plots also highlight alternating N–S strips of negative and positive values (± 15 cm), clearest in the northern half of the survey. We interpret these as errors in the ULS point cloud due to a poorly calibrated INS on some of our flights.

6 The Eastern Denali fault: maximal coverage of a major strike-slip fault

6.1 Background and motivations

The Denali fault hosted North America's largest and longest onshore earthquake of the modern instrumental period. The M_w 7.9 earthquake of November 3 2002 ruptured for $\sim$340 km west to east across central Alaska, producing mostly right-lateral surface offsets of up to $\sim$9 m (Eberhart-Phillips et al., 2003). Though the Denali fault continues southeastwards into Yukon and northwestern BC (Fig. 2A), the 2002 earthquake stopped short of the Canadian border, branching instead onto the Totshunda splay fault, where it terminated. This rupture pattern has elicited investigations into the current activity and kinematics of the Denali fault east of the Totshunda junction (Bostock, 1952; Clague, 1975; Haeussler et al., 2017; Marechal et al., 2018; Blais-Stevens et al., 2020; Choi et al., 2021), usually referred to as the Eastern Denali fault (EDF). The EDF has been active in the Holocene since it lacks a glacial overprint and displays several push-up or mole-track structures within till (Bostock, 1952; Blais-Stevens et al., 2020). Paleoseismic trenching of the fault and coring of lake sediments ponded against the scarp also revealed evidence for five strong earthquakes during the past $\sim$6,800 years, leading Blais-Stevens et al. (2020) to call for the acquisition of lidar imagery to better illuminate the surface offsets, kinematic style, and other characteristics of these events.

We targeted a $\sim$100 km stretch of the EDF centered upon Lù'àn Män (Kluane Lake) and paralleling the paved Alaska Highway (Fig. 11D). This section of the fault occupies a broad glacial valley, surfacing up to a few kilometers NE of the frontal range of the St. Elias mountains and displays tectonic landforms including those targeted for paleoseismic trenching and coring by Blais-Stevens et al. (2020). The area is mostly covered by boreal forest (Fig. 2A), consisting mainly of white spruce trees, aspen and balsam poplar. However, the EDF crosses several wide fluvial terraces deposited by rivers sourced in the St. Elias mountains, the youngest of which are only sparsely vegetated. Our work along the EDF therefore provides an example of surveying rugged (though not mountainous) topography containing both dense and sparse vegetation. Our surveying of the EDF also represents our closest attempt at a regional ULS survey. We sought to survey as much of the fault as possible, but sparse secondary road coverage off the main Alaska Highway prevented us from accessing long stretches of it. We therefore flew several separate sections of the EDF from launch sites located wherever a passable road crosses the fault, usually one that follows a major river sourced in the St. Elias mountains (Fig. 11D). However, these river outlets are also where we expect to observe some of the best expressions of the fault in Quaternary deposits, such as deformed river terraces and offset terrace risers. Our drone coverage, though discontinuous, should therefore still capture many of the features of greatest geomorphic interest.

6.2 ULS data acquisition and results

We collected $\sim$10 km^2 of drone lidar data at seven individual survey sites—from NW to SE, Quill Creek, Burwash Creek, Duke River, Copper Joe Creek, Nines Creek, Slims River/Topham Creek, and Telluride Creek—that together capture a $\sim$15 km length of the EDF (Fig. 11D).

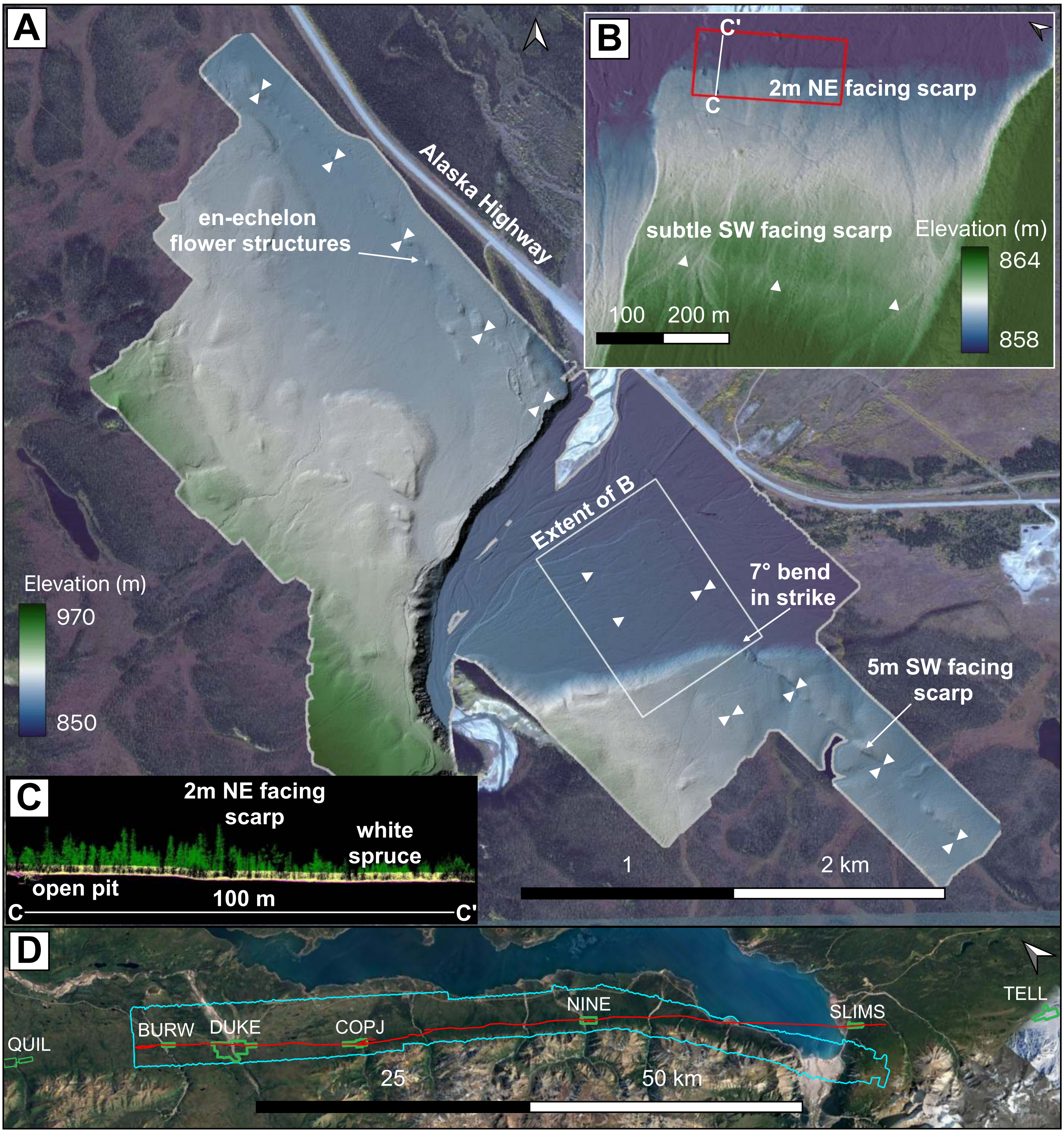


Figure 11 (A) Hillshaded drone lidar DTM of the Duke River site on the Eastern Denali fault (EDF). The extent of panel B is marked by the white rectangle. White triangles indicate the approximate location of the EDF. Underlying satellite imagery is from Bing. (B) Inset showing a 2 m high northeast facing scarp along a river terrace. The red polygon shows the bounds of the flight parameter testing discussed in Section 7.1. White line shows the location of cross section C–C'. (C) Cross section C–C' through the classified ULS point cloud, with green for vegetation, pink for ground returns, and yellow for unclassified points. (D) Locations of all ULS collection sites (green polygons) along the trace of the EDF as identified with lidar data (red line). The cyan polygon outlines the comparison ALS dataset. Satellite imagery is from Google. QUIL = Quill Creek, BURW = Burwash Creek, DUKE = Duke River, COPJ = Copper Joe Creek, NINE = Nines Creek, SLIMs = Slims River/Topham Creek, TELL = Telluride Creek.

Data were acquired during two, week-long field campaigns in September 2021 and August 2022, each involving three crew members (a pilot and two visual observers); the largest site, at Duke River, was flown over multiple days (Fig. 11A). Flight paths for each site are plotted in supplementary figures SM5–SM11. The drone was flown at 80 m AGL and speeds of 5 m s^{-1} in 2021 and 4 m s^{-1} in 2022. Since large portions of each target area were difficult to access on foot, we were unable to deploy as many GCPs as we did at the study sites described in previous sections. The number of GCPs range from 2 at the Nines and Quill Creek sites to 20 at the Duke River site (cumulative across several days of surveying).

The seven ULS surveys yielded average point densities of 80–112 pts/m^2 with a mean of 98 pts/m^2, average classified ground return densities of 20–72 pts/m^2 with a mean of 45 pts/m^2, and average ground return spacings of 0.19–0.23 m with a mean of 0.21 m (see Table 3 for results of each individual survey). Generally, the ULS does an excellent job of imaging beneath the boreal forest canopy (inset, Fig. 11A), with ∼47% of all laser returns classified as ground. We did encounter some misalignment of flight lines in some of the datasets—particularly at Quill Creek—again, potentially due to a poor INS calibration. This can produce some striping in the M3C2 and DoD results (Fig. 12), with separations of ±15 cm in places. However, these linear artefacts are easily distinguished from genuine tectonic landforms. The ground returns were rasterized at an optimal pixel resolution of 0.3 m for each of the seven individual surveys.

We focus our further analysis in this paper on the Duke River site, which is the largest and densest of our EDF ULS surveys (Fig. 11 and Table 3). Results from the other EDF drone lidar surveys are presented in the supplemental material (SM14—16); some of them were also interpreted in an earlier technical report of ours (Finley et al., 2022a) and a full tectonic analysis of all of the EDF lidar topography will be the subject of a future paper. The 0.3 m-resolution Duke River DTM showcases several interesting tectonic landforms along the principal trace of the EDF (Fig. 11A). From NW to SE, these include en-echelon push up structures indicative of dextral strike-slip, a pair of clear, right-lateral offsets to terrace risers south of the active Duke River, an abrupt, 7° bend in fault strike, and a large, SW-facing scarp with a vertical separation of ∼5 m. The lidar DTM also reveals a previously unmapped, secondary strand of the EDF, expressed as a SW-facing scarp crossing the widest, southern terrace of the Duke River (Fig. 11B). This highlights the potential for drone lidar to capture subtle, off-fault tectonic landforms away from principal fault traces.

6.3 Comparison and differencing with ALS data

Until this study, the Canadian portion of the Denali fault had not been flown systematically with lidar, with the best freely-available topographic data coverage being the 2 m-resolution, satellite photogrammetry-derived ArcticDEM (Morin et al., 2016) and a bespoke 4 m-resolution DTM constructed from legacy airphotos using SfM (Bender and Haeussler, 2017). In addition to our drone lidar surveying described above, one of us (B.M.) also collected a ∼295 km^2, ∼70 km-long airborne lidar swath on 19 August 2018 that captures ∼50 km of the EDF between Burwash Creek and south of Nines Creek (Fig. 11D), which for the purposes of this study we use as our comparison dataset. The ALS survey covers four of our seven drone lidar sites, including the largest at Duke River. The airborne lidar yields an average point density of ∼8 pts/m^2 and an average ground return density of ∼3.5 pts/m^2, less than 10% of the equivalent ULS values, and the average ALS ground return spacing is 0.54 m, more than double that of the ULS (Table 3). Consequently, the ULS DTM exhibits noticeably finer detail than is discernible in the ALS DTM.

For the Duke River dataset, the ICP rigid body transformation that optimally aligns the ULS and ALS point clouds has a translation vector of 0.54 m and a maximum rotation of 0.0003°, indicating global registration to within a few decimeters and negligible tilting (Table 3). The mean M3C2 distance between the aligned point clouds is 0.00 m with a standard deviation of 0.18 m, while the DoD has an average elevation difference of 0.00 m with a standard deviation of 0.28 m, indicating excellent internal consistency between the two lidar surveys. The other survey sites with paired coverage have slightly higher global ICP translations (of up to 0.99 m), but had similarly low mean M3C2 and DoD values (of up to 0.01 m) and standard deviations (of up to 0.28 m).

After ICP alignment of the ULS and ALS ground returns, the largest M3C2 distances between the two clouds are along the active braided channels of the Duke River, with some showing erosion and others deposition such as from the formation of sandbars (Fig. 12A). There are also some large M3C2 distances found along the bluffs on the northern bank of the Duke River. These steep slopes are mostly composed of unconsolidated glacial till and are very unstable, with several small failures occurring while we were working in the area. Thus, it is unsurprising that there are differences in these surfaces between the ALS and ULS acquisitions. Other areas with non-negligible M3C2 values may result from slight misaligments between the individual, day-to-day acquisitions at the Duke River site, which were georeferenced using separately-surveyed GCP deployments. The raster DoD highlights similar areas of difference (Fig. 12B), although it does appear a little noisier than the M3C2 distances due to small misalignments in the rasterization of roads, small channels and other linear features.

7 Discussion

7.1 Flight parameter trade-offs

The four case studies described in Sections 3–6 highlight not only the usual trade-off in remote sensing between data spatial resolution and coverage, but also the flexibility in drone surveying to adjust flight parameters for the job at hand in a way that would be difficult with a crewed aircraft platform. For example, for our local survey of the eastern X̲EOLX̲ELEK (Elk Lake) shoreline site, we flew the drone at a substantially slower speed and a lower height than in the other, kilometric-scale surveys, resulting in a point cloud that was around five times denser but also smaller in scale (Table 3). In practice, however, the trade-off between platform speed and point cloud density is complicated by the limited battery life of the drone, which restricts the area that can be collected in a single flight.

We therefore conducted an explicit test of the trade-

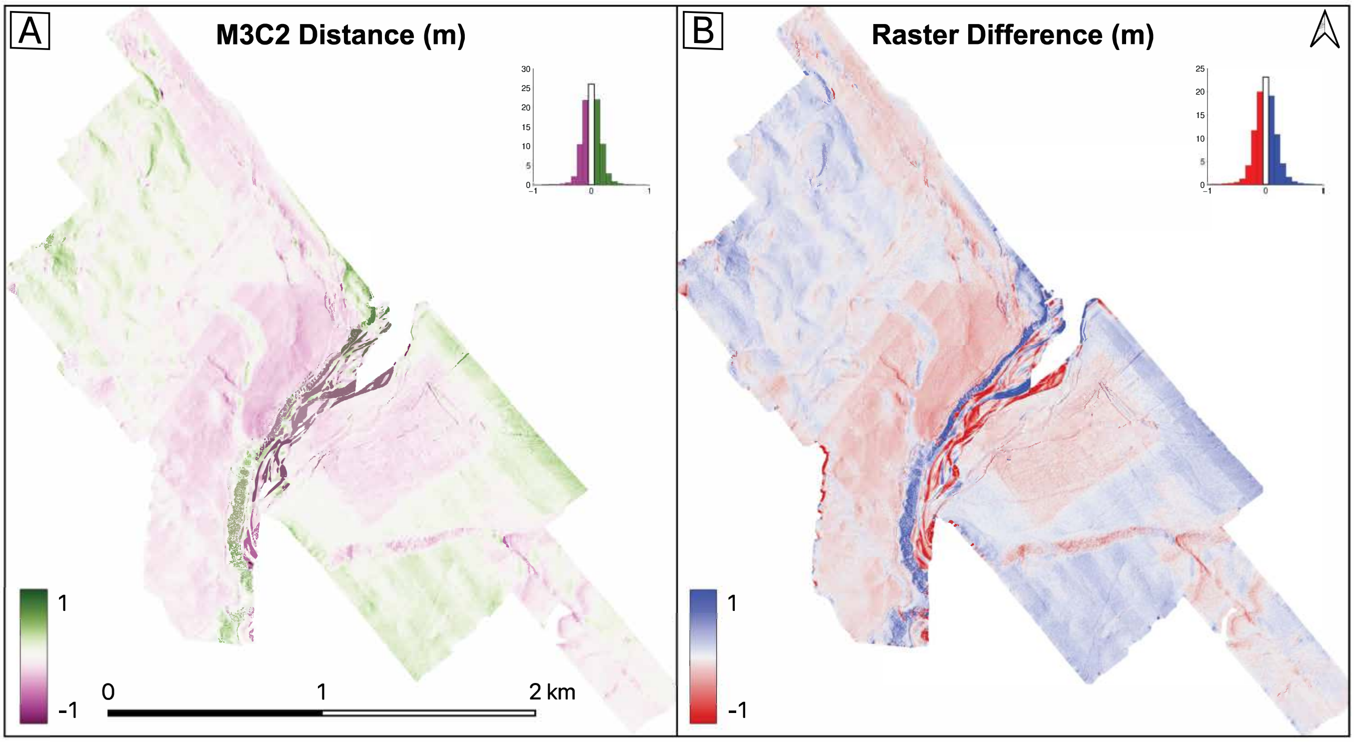


Figure 12 Differencing of the drone and airborne lidar datasets for the Duke River site along the EDF, shown as (A) M3C2 distances calculated in CloudCompare and (B) a DEM of Difference. Positive values indicate where the ULS dataset was higher than the ALS dataset. Histograms show distributions of raster values.

offs between platform speed, flight duration, and lidar point density, with the aim of determining the ideal flight parameters for collecting high-density data as efficiently as possible. We did so during our surveying along the EDF in 2022, choosing for the test a small ($\sim$2500m^2) area of the Duke River site centered along a segment of the EDF that offsets an abandoned, forested river terrace (Fig. 11B). We surveyed this area nine times over the course of a single day, at 1 m s^{-1} increments in speed from 1 m s^{-1} to 10 m s^{-1}, using the same optimal flight height (80 m AGL) and the same flight pattern (Fig. 13B, left panel). The nine flights yielded order of magnitude ranges in both point densities, from $\sim$51–496 pts/m^2, and classified ground return densities, from $\sim$10–68 pts/m^2 (Table 4). We observe the expected tradeoff between platform speed and ground return density, with the slowest flight yielding more than double the point density of the next slowest flight but taking double the time (Fig. 13A). The resulting DTMs all looked similar to the eye (Fig. 13B), although the data from faster collections were slightly noisier. The faster collections had a greater variation in the DTM surface, potentially as a result of small bushes and other low vegetation being classified as ground.

Ultimately, the most efficient speed for a given survey depends on the desired point density and should be determined on a case-by-case basis accounting for both the scale of the features being targeted and the vegetation cover type. In our general case, we strive for submeter pixel DTMs in order to identify fine-scale fault geomorphology that might not be visible in existing airborne datasets. Ideally, each raster cell value should be based on the average of at least 3 ground returns. Thus, for a 50 cm DTM, 12 ground returns per square meter are desirable, allowing each raster cell to be calculated using a minimum of 3 ground points. The minimum DTM resolution in Table 4 was calculated using this rule of thumb. Using Fig. 13A as a guide, an appropriate maximum acquisition speed would therefore be about 6 m s^{-1}. Anecdotally, this agrees well with our experience gleaned from many drone lidar campaigns.

7.2 Drone lidar performance

Our four case studies described in Sections 3–6, as well as the additional testing at the Duke River site described above, demonstrate the wide range of ground return densities attainable with drone lidar, governed principally by the platform height and speed, swath overlap, and vegetation type. At its densest—our local survey of a planned paleoseismic site along the XELF with only scattered tree cover (Section 3)—we obtained a ground point density of 260 pts/m^2 at an average spacing of 6 cm, though of course further improvements would have been possible with additional, overlapping flight lines. In surveys undertaken at the kilometric length scales more generally of interest to tectonic geomorphologists, we obtained ground return densities ranging from $\sim$10 pts/m^2 (spacing of $\sim$30 cm) along the rugged, heavily forested SJF (Section 4) to $\sim$70 pts/m^2 (spacing of $\sim$20 cm) at the mixed-cover Duke River site on the EDF (Section 6). In all cases, the ULS densities were a marked improvement from those of the comparison ALS datasets, which ranged from 3.5–9 pts/m^2 with spacings of 0.33-0.54 m (Table 3). Of course, these differences trade off against aerial coverage; in our

Speed (*m s*$^{-1}$)	*Survey time* (*min*)	*Point density* (*pts/m*^{2})	*Ground point density* (*pts/m*^{2})	*%pts ground*	*Minimum DTM resolution* (*m*)
1	23.50	495.63	68.36	14	0.21
2	11.75	242.49	28.16	12	0.33
3	7.83	155.68	23.80	15	0.36
4	5.88	122.87	20.10	16	0.39
5	5.37	99.83	19.20	19	0.40
6	4.47	84.98	15.58	18	0.44
7	3.83	72.70	12.12	17	0.50
8	3.35	65.17	14.19	22	0.46
9	2.98	57.93	9.70	17	0.56
10	2.68	50.52	10.05	20	0.55

Table 4 Results of our testing of platform speed trade-offs. The survey time does not include the static IMU calibrations (∼5 minutes per flight), nor the time taken to transit the drone between the launch site and the start of the first data collection flight line. The minimum DTM resolution is based on a recommendation that at least 3 ground points should be averaged per raster cell.

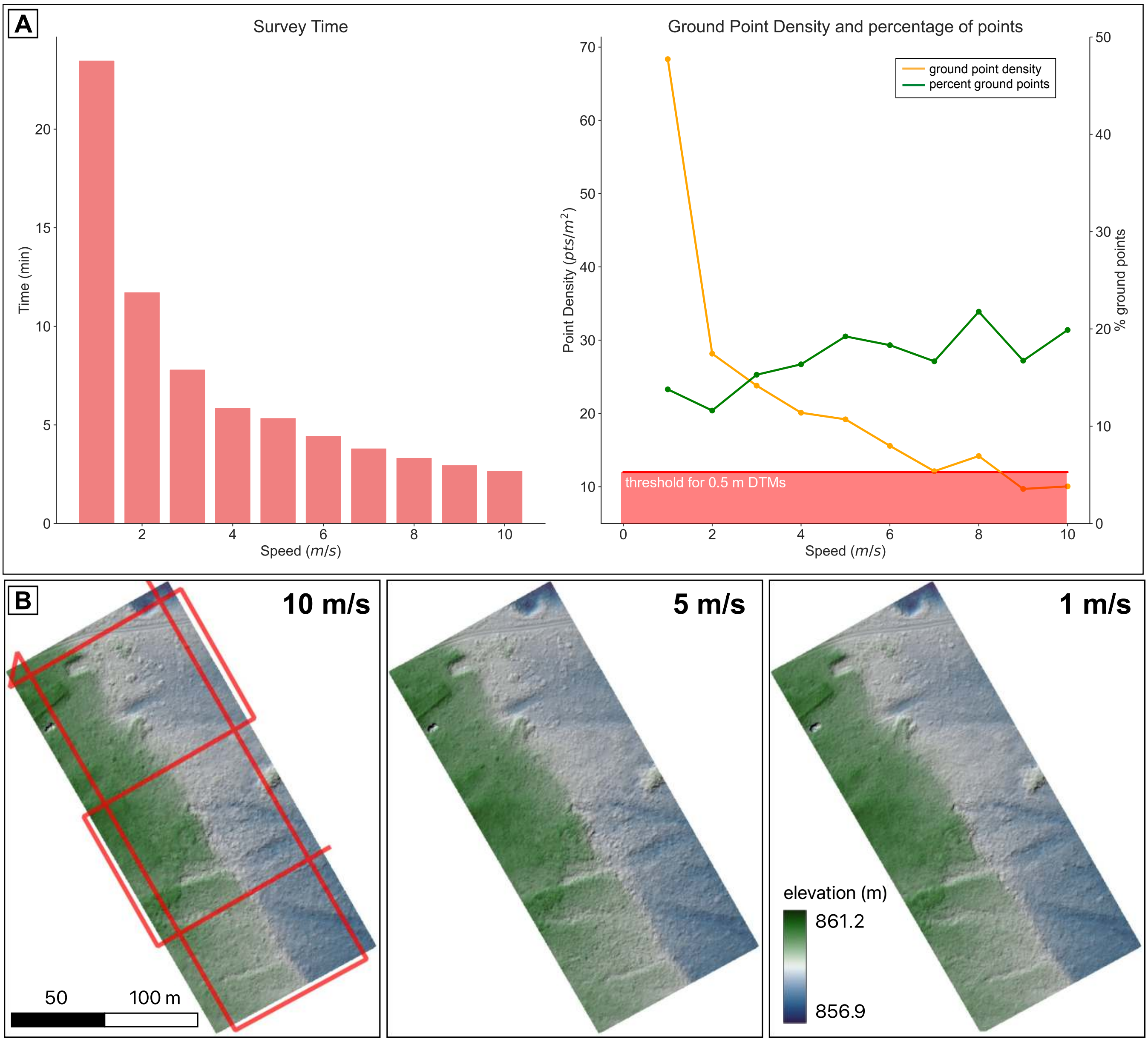


Figure 13 Results of platform speed trade off tests. (A) Relationships between acquisition time, platform speed, and the resulting ground return density. The 12 points per square meter threshold illustrates that sub-0.5 cm DTMs can only reliably be obtained at speeds of 6 m s^{-1} or less. (B) 30 cm-resolution DTMs for the fastest (10 m s^{-1}), intermediate (5 m s^{-1}) and slowest (1 m s^{-1}) flights. The red line on the 10 m s^{-1} panel shows the acquisition path used for all flights.

largest ULS survey along the EDF we collected a total of ∼10 km^2 of data in two week-long field campaigns, whereas our ALS survey collected 25 times that area in a single day. Our ULS surveying along the SJF, SRTM and EDF shows that coverage of 0.5–1.6 km^2 is achievable in a single day with the drone, including over rugged topography.

Our ICP alignments of ULS ground returns with corresponding ALS point clouds show that the two datasets are usually globally registered to within less than a meter of one another (Table 3), with the one exception, along the SJF, likely arising from differences in the geoid model used to calculate orthometric heights (Section 4.3). The Riegl MiniVUX-1UAV laser scanner has expected accuracy of 10 mm and precision of 5 mm and the Applanix APX20-UAV INS has a post-processed accuracy of 2–5 cm. Once the ULS and ALS ground return clouds are aligned using ICP, average M3C2 distances and DoD values on the order of a few centimeters are therefore within the expected noise, especially considering that the average is biased by localized occurrences of significant landscape alteration between surveys (e.g. road construction; Fig. 10B–C). Generally, the M3C2 distances were very similar to the DoD values, implying that the local distances calculated on the point clouds were mainly in the vertical direction. We did encounter artefacts arising from misaligned flight lines within some of the ULS datasets (e.g. Figs. 7B and 9), although further post-processing may have helped reduce these (Gu et al., 2023) and similar problems can occur in some ALS surveys, too (Scott et al., 2022). One main challenge for mobile lidar systems thus far has been in point accuracy, particularly as a result of the IMU trajectory Glennie et al. (2013), and our ULS results bear this out.

Our detailed comparisons of data collected at the eastern X̱EOLX̱ELEꝀ (Elk Lake) shoreline site show how ULS penetrates low-lying vegetation (blackberry bushes) better than ALS does, with both lidar systems naturally outperforming SfM (Section 3.3). Counterintuitiuvely, the ULS produces a lower fraction of ground returns (48%) than the ALS (64%), which hints that in densely-vegetated areas, some laser returns off shrubs and bushes might be misclassified as ground in ALS datasets, to a greater extent than they are in ULS datasets. This would explain why the ground surface modelled from ULS data is often slightly lower than that from ALS data (Figs. 7 and 9). It is important to note that many ground classification algorithms, such as progressive morphological filters (Zhang et al., 2003) and cloth simulation functions (Zhang et al., 2016), were developed using and for ALS data. Thus to effectively determine ground points within ULS data, the default algorithm parameters may need to be adjusted.

The relative vegetation penetration capabilities of ULS and ALS are further showcased in Figure 14. In the left-hand column, we plot 5 m wide point cloud swath cross-sections for typical vegetation present in each of the study sites. These demonstrates that the ULS is better at both capturing the vegetative structure and minimizing gaps in the ground returns. This may partly reflect that the ULS footprint is ∼6–9 cm in diameter at our optimal flight elevation of 80 m AGL, several times narrower than the ∼15–90 cm footprints typical of airborne systems (e.g. Lin et al., 2013; Fernandez-Diaz et al., 2014).

Tectonic geomorphologists generally analyse rasterized DTMs rather than point clouds and so it is important to consider ULS in this context. The ∼1–10 pts/m^2 ground return densities typical of ALS data (Figure 14, middle column) usually translate to ∼1 m raster resolutions, whereas the much the denser point clouds collected by our drone system (Figure 14, right-hand columns) allow for finer pixel dimensions of 0.2–0.5 m. Lin et al. (2013) proposed that 0.5 m was an optimal DTM resolution for detecting tectonic-geomorphic signals, with 0.25 cm pixels offering little improvement, but this may reflect the larger (0.41 m) footprint of their airborne laser scanner. Since ULS laser footprints are much smaller (6–9 cm), it enables us to produce sub-50 cm DTMs which aid the identification of many features obscured in typical 1 m resolution imagery, such as ruts in road-ways, footpaths, and individual tree stumps or logs (e.g. Figs. 7 and 8B).

7.3 Limitations and future prospects of drone lidar

Based on our four case studies in Sections 3–6 and the performance metrics discussed above, we foresee a number of specific applications for drone lidar within the field of tectonic geomorphology. Because battery life, road access, and VLOS requirements limit us to kilometric fault length-scales (e.g. Fig. 11D), we do not envisage ULS (or any other type of drone-based imaging) as a regional reconnaissance tool in the way that ALS has become. However, ULS may be a useful, relatively low-cost way of extending lidar coverage beyond the footprint of an existing ALS survey, such as we did along both the SRMT and EDF (Figs. 10A and 11D). As demonstrated in Section 3 and 7.2, ULS also offers better and more even ground point coverage beneath trees and shrubs than ALS, making it possible to densify lidar coverage along known faults in vegetated landscapes. Faulted landforms targeted for paleoseismic trenching, slip rate studies, or morphologic dating may benefit from the finer, decimetric spatial resolutions achievable with ULS. We have also shown how drone lidar can reveal subtle landforms associated with structural complexity and distributed deformation, such as the secondary scarps imaged on the Columbia Lake alluvial fans (Fig. 10D) and at the Duke River site along the EDF (Fig. 11). There is also scope for mapping landscapes and landforms associated with other natural hazards, such as landslides (Pellicani et al., 2019), volcanoes, tsunamis, and flooding. Finally, because drone deployments are both logistically easier and cheaper than procuring a crewed aircraft, there is rich potential for the use of ULS for rapid response (e.g. mapping of new earthquake surface ruptures) and in repeat mode (e.g. high temporal and spatial resolution time-series of postseismic afterslip). Building upon this study, the capability of repeat ULS surveying for mapping three-dimensional surface deformation will be tackled explic-

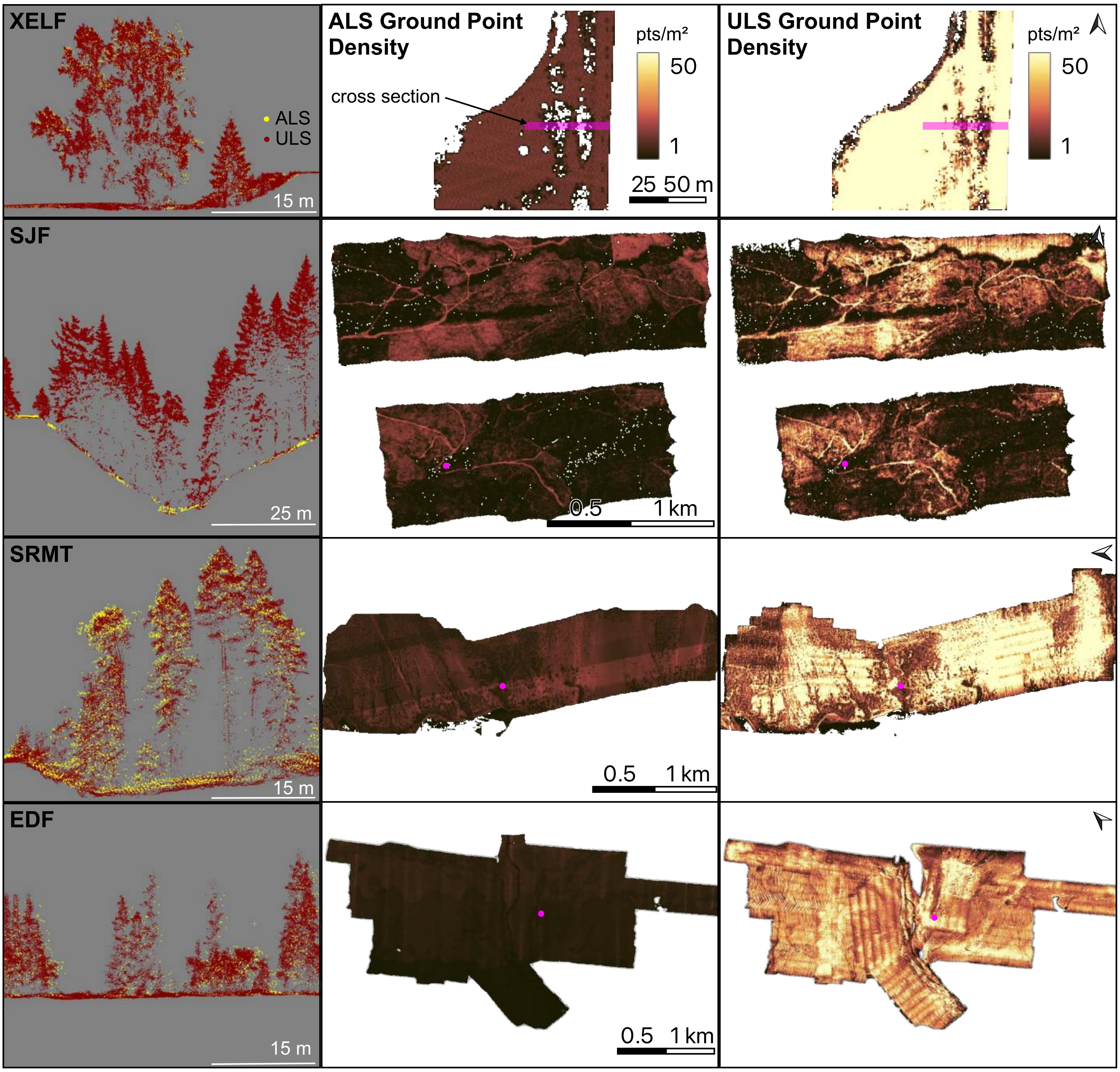


Figure 14 Left-hand column: representative 5 m-wide swath cross-sections through the ULS (red) and ALS (yellow) point clouds showing typical vegetation for the XELF (top row), SJF (second row), SRMT (third row) and EDF (bottom row) study sites. The point clouds are displayed concurrently, with ALS points enlarged to prevent them from being obscured behind ULS points. Note that for the SJF site, only ALS ground returns were provided to us. Middle column: ground point density maps for the ALS datasets at the four study sites. Pink polygons mark the cross-section locations of the left-hand column. Right-hand column: ground point density maps for the ULS datasets. Additional ULS point cloud profiles through stereotypical vegetation for each survey site are provided in supplemental figures SM17—20.

itly in a forthcoming paper of ours.

Just as ULS will complement rather than supplant ALS as a source of high-resolution topographic data, we do not see it replacing TLS or drone-based SfM systems for all applications. Firstly, the upfront costs of our ULS system (discussed further below) are higher than that of many TLS systems and much higher than those of drone-based SfM. Since many consumer drones come equipped with high-quality cameras (e.g. Pellicani et al., 2019), we envisage that SfM will remain the tool of choice for collecting high-resolution topography in arid, vegetation-sparse landscapes (e.g. Johnson et al., 2014). ULS would still offer certain advantages over SfM in such regions—for example its active source allows mapping in low light conditions, and the processing of the lidar point cloud is significantly less computationally intensive—but these would come at greatly increased cost. Similarly, ULS is unlikely to replace TLS for applications in which ground-based scanner vantage points suffice. However, ULS can capture a small target area quicker than TLS; for example, Brede et al. (2017) collected a 140 pts/m^2 ULS dataset in a single 9 minute flight, whereas it took two days to complete TLS coverage of the same 100 m $\times$ 180 m site.

The components of our ULS system cost more than CAD $100,000 when purchased in 2018–2019, though these costs are likely to decrease significantly as the technology matures (e.g. Van Tassel, 2021). As ULS is adopted for a broader range of scientific and commercial applications, an increasing variety of drones and miniaturized scanners and INS instruments are becoming available, including as part of integrated systems. In our experience, once the upfront costs of purchasing a ULS system is made, the year-on-year costs for insurance and software licensing are inexpensive. One of the biggest constraints on drone systems is their limited flight time, related to battery life. However, electric cars and a growing array of battery powered mobile devices are driving improvements in battery technology that will positively benefit drone surveying (Townsend et al., 2020; Liang et al., 2019; Rajashekara, 2013, e.g.). Alternatively, gasoline and hybrid-powered UAVs are becoming commercially available, offering much higher energy densities and ensuing improvements in both flight time and lifting capacity (e.g. Skyfront, 2023; Metcalf et al., 2022; Viswanathan et al., 2022; Harris Aerial, 2023). Another major constraint for ULS is imposed by aviation authorities, which require specialist equipment and large amounts of permitting for beyond visual line-of-sight (BVLOS) flight in both Canada and the U.S. Flying BVLOS would allow for larger drone acquisitions with fewer crew and less time in the field. This will potentially change as drone software and collision avoidance systems mature and new RPAS-specific legislation emerges (Transport Canada, 2023).

8 Conclusions

We describe a state-of-the-art drone lidar system, provide a practical guide for other researchers interested in developing their own, and showcase its performance using four case studies from a range of terrain and vegetation types found within the Canadian Cordillera. These range from a local ($\sim$100 m $\times$ 100 m) survey of a paleoseismic trenching site with scattered tree cover, captured at a ground return density of 260 pts/m^2, to multi-kilometer mapping of faults in remote, forested regions, captured with ground return densities of $\sim$10–72 pts/m^2. Our ULS point clouds are gridded into bare earth DTMs with $\sim$20–50 cm pixel dimensions, substantially finer than the $\sim$1 m dimensions typical of airborne lidar DTMs. In most cases, the drone lidar ground returns are globally registered to overlapping airborne lidar data to within $\sim$0.3–1.0 m, and once aligned, point-to-point distances and DEMs of difference indicate internal consistency to within a few centimeters. Distinct advantages of terrain mapping using ULS include better imaging beneath vegetation, the flexibility to adjust flight parameters to achieve a desired ground return density, and relatively straightforward platform deployment logistics. In practice, ULS mapping is currently limited to kilometric lengthscales by battery life, road access requirements, and regulatory constraints, so it is unlikely to replace ALS for regional fault reconnaissance. However, we envisage rich potential of drone lidar for (1) cost-effectively mapping faults beyond the edges of existing ALS surveys; (2) detailed surveying of known faults for paleoseismic trenching, fault slip rate estimations, or morphologic dating; (3) revealing subtle landforms arising from off-fault deformation; (4) rapid collection of perishable data such as along earthquake surface ruptures; and (5) for repeat deployments along surface ruptures for capturing afterslip.

Acknowledgements

We are grateful to have conducted this work on the traditional territories of the W̱SÁNEĆ people (XELF), the Hul'q'umi'num speaking people (SJF), the Ktunaxa and Secwépemc (SRMT), and the Lù'àn Män Dän (Kluane Lake People, EDF). Our ULS system was supported through the Canada Foundation for Innovation John Evans Leadership Fund (CFI JELF) and the Government of British Columbia Knowledge Development Fund (BCKDF), and we also acknowledge generous educational discounts and in-kind contributions from Riegl, Applanix, Candrone, and Cansel. Fieldwork in BC was primarily supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) through Discovery Grants #2017-04029 and #2019-05639 (to E.N. and B.M. respectively) and fieldwork in the Yukon by the Yukon Geological Survey. G.S. was supported in part through an Agnew Memorial Scholarship from the University of Victoria, T.F. through an NSERC Alexander Graham Bell Canada Graduate Scholarship, and E.N. and B.M. through Canada Research Chairs. We thank Chris Bone at University of Victoria's Department of Geography for sharing lab space, equipment, and running costs so generously. We are grateful to Chris Crosby for facilitating the hosting of our drone lidar data on the OpenTopography web portal, which is supported by the National Science Foundation under NSF Award Numbers 1948997, 1948994 and 1948857. We thank the Capital Region District (in particular April Mitchell, Jeannette Mollin and Marc Solomon) for allowing us to work within the Elk/Beaver Lake Regional Park, BC Parks for granting us access to Columbia Lake Park, the Yukon Geological Survey (in particular Maurice Colpron, Carolyn Relf and Sarah Sternbergh) and the Arctic Institute of North America's Kluane Lake Research Station for facilitating fieldwork along the EDF, and Mosaic Forest Management for sharing airborne lidar data along the SJF and for granting us access to field sites there. We also thank our University of Victoria colleagues Isaraporn Sethanant, Élyse Gaudreau, Jinrui Liu, and Braeden Clark for their assistance as visual observers during our drone surveys, as well as Andrew Schaeffer, Lucinda Leonard, Eva Kwoll and John Cassidy for their suggestions and support of the project. Finally, we are grateful to Nadine Reitman, Randy Williams, and an anonymous reviewer for their thoughtful comments that improved the manuscript.

Data and code availability

Supplemental material can be found on Zenodo (https://doi.org/10.5281/zenodo.10092585). This includes shell scripts for the classification and rasterization of point cloud data collected by a UAV laser scanning (ULS) system. The Zenodo folder also includes visualizations of flight paths for each survey and data comparisons made at the BURW, COPJ and NINE sites along the Eastern Denali fault that were not included within the publication, as well as ULS point cloud cross sections that showcase stereotypical vegetation for each case study. All drone lidar data are available on the OpenTopography portal at the following links: XELF (https://doi.org/10.5069/G92V2DBC), SJF (https://doi.org/10.5069/G9TB1542), SRMT (https://doi.org/10.5069/G9PK0DCW), EDF (https://doi.org/10.5069/G998857Q). Airborne lidar data for the XELF (Section 3) and SRMT (Section 5) sites are available on the LidarBC (2023) portal.

Competing interests

The authors have no competing interests.

References

Arrowsmith, J. R., Rhodes, D. D., and Pollard, D. D. Morphologic dating of scarps formed by repeated slip events along the San Andreas Fault, Carrizo Plain, California. Journal of Geophysical Research (Solid Earth), 103(B5):10,141–10,160, 1998. doi: 10.1029/98JB00505.

Baldwin, K., Allen, L., Basquill, S., Chapman, K., Downing, D., Flynn, N., MacKenzie, W., Major, M., Meades, W., Meidinger, D., Morneau, C., Saucier, J.-P., Thorpe, J., and Uhlig, P. Vegetation Zones of Canada: A Biogeoclimate Perspective, 2019.

Bemis, S. P., Micklethwaite, S., Turner, D., James, M. R., Akciz, S., Thiele, S. T., and Bangash, H. A. Ground-based and UAV-Based photogrammetry: A multi-scale, high-resolution mapping tool for structural geology and paleoseismology. Journal of Structural Geology, 69:163–178, 2014. doi: 10.1016/j.jsg.2014.10.007.

Benavente, C., Wimpenny, S., Rosell, L., Robert, X., Palomino, A., Audin, L., Aguirre, E., and García, B. Paleoseismic Evidence of an M_w 7 Pre-Hispanic Earthquake in the Peruvian Forearc. Tectonics, 40(6):e2020TC006479, 2021. doi: 10.1029/2020TC006479.

Bender, A. M. and Haeussler, P. J. Eastern Denali Fault surface trace map, eastern Alaska and Yukon, Canada. USGS Numbered Series 2017-1049, U.S. Geological Survey, Reston, VA, 2017.

Bertiger, W., Bar-Sever, Y., Dorsey, A., Haines, B., Harvey, N., Hemberger, D., Heflin, M., Lu, W., Miller, M., Moore, A. W., Murphy, D., Ries, P., Romans, L., Sibois, A., Sibthorpe, A., Szilagyi, B., Vallisneri, M., and Willis, P. GipsyX/RTGx, a new tool set for space geodetic operations and research. Advances in Space Research, 66(3):469–489, 2020. doi: 10.1016/j.asr.2020.04.015.

Blais-Stevens, A., Clague, J. J., Brahney, J., Lipovsky, P., Haeussler, P. J., and Menounos, B. Evidence for Large Holocene Earthquakes along the Denali Fault in Southwest Yukon, Canada. Environmental & Engineering Geoscience, 26(2):149–166, 2020. doi: 10.2113/EEG-2263.

Bostock, H. S. Geology of northwest Shakwak Valley, Yukon Territory. E. Cloutier, Queen's Printer, 1952.

Brandon, M. T. Origin of igneous rocks associated with melanges of the Pacific Rim Complex, western Vancouver Island, Canada. Tectonics, 8(6):1115–1136, 1989. doi: 10.1029/TC008i006p01115.

Brede, B., Lau, A., Bartholomeus, H. M., and Kooistra, L. Comparing RIEGL RiCOPTER UAV LiDAR Derived Canopy Height and DBH with Terrestrial LiDAR. Sensors, 17(10), 2017. doi: 10.3390/s17102371.

Brooks, B. A., Glennie, C., Hudnut, K. W., Ericksen, T., and Hauser, D. Mobile Laser Scanning Applied to the Earth Sciences. Eos, Transactions American Geophysical Union, 94(36): 313–315, 2013. doi: 10.1002/2013EO360002.

Brooks, B. A., Minson, S. E., Glennie, C. L., Nevitt, J. M., Dawson, T., Rubin, R., Ericksen, T. L., Lockner, D., Hudnut, K., Langenheim, V., Lutz, A., Mareschal, M., Murray, J., Schwartz, D., and Zaccone, D. Buried shallow fault slip from the South Napa earthquake revealed by near-field geodesy. Science Advances, 3(7):e1700525, 2017. doi: 10.1126/sciadv.1700525.

Bubeck, A., Wilkinson, M., Roberts, G. P., Cowie, P. A., McCaffrey, K. J. W., Phillips, R., and Sammonds, P. The tectonic geomorphology of bedrock scarps on active normal faults in the Italian Apennines mapped using combined ground penetrating radar and terrestrial laser scanning. Geomorphology, 237: 38–51, 2015. doi: 10.1016/j.geomorph.2014.03.011.

Butler, H., Chambers, B., Hartzell, P., and Glennie, C. PDAL: An open source library for the processing and analysis of point clouds. Computers & Geosciences, 148:104680, 2021. doi: 10.1016/j.cageo.2020.104680.

Chen, T., Akciz, S. O., Hudnut, K. W., Zhang, D. Z., and Stock, J. M. Fault-Slip Distribution of the 1999 Mw 7.1 Hector Mine Earthquake, California, Estimated from Postearthquake Airborne LiDAR Data. Bulletin of the Seismological Society of America, 105 (2A):776–790, 2015. doi: 10.1785/0120130108.

Choi, M., Eaton, D. W., and Enkelmann, E. Is the Eastern Denali fault still active? Geology, 49(6):662–666, 2021. doi: 10.1130/G48461.1.

Civil Aviation Safety Authority. Remote pilot licence, Mar. 2021. https://www.casa.gov.au/drones/get-your-operator-credentials/remote-pilot-licence.

Clague, J. J. Late Quaternary Sediments and Geomorphic History of the Southern Rocky Mountain Trench, British Columbia. Canadian Journal of Earth Sciences, 12(4):595–605, 1975. doi: 10.1139/e75-054.

Clark, K. J., Nissen, E. K., Howarth, J. D., Hamling, I. J., Mountjoy,

J. J., Ries, W. F., Jones, K., Goldstien, S., Cochran, U. A., Villamor, P., Hreinsdóttir, S., Litchfield, N. J., Mueller, C., Berryman, K. R., and Strong, D. T. Highly variable coastal deformation in the 2016 M_W7.8 Kaikōura earthquake reflects rupture complexity along a transpressional plate boundary. Earth and Planetary Science Letters, 474:334–344, 2017. doi: 10.1016/j.epsl.2017.06.048.

Cui, Y., Miller, D., Schiarizza, P., and Diakow, L. British Columbia digital geology. British Columbia Ministry of Energy, Mines and Petroleum Resources, British Columbia Geological Survey Open File, 8(9), 2017. Data version 2019-12-19.

Cunningham, D., Grebby, S., Tansey, K., Gosar, A., and Kastelic, V. Application of airborne LiDAR to mapping seismogenic faults in forested mountainous terrain, southeastern Alps, Slovenia. Geophysical Research Letters, 33(20):L20308, 2006. doi: 10.1029/2006GL027014.

Căţeanu, M., Arcadie, C., and others. ALS for terrain mapping in forest environments: An analysis of LiDAR filtering algorithms. EARSeL eProceedings, 16(1):9–20, 2017. doi: 10.12760/01-2017-1-02.

DeLong, S. B., Lienkaemper, J. J., Pickering, A. J., and Avdievitch, N. N. Rates and patterns of surface deformation from laser scanning following the South Napa earthquake, California. Geosphere, 11(6):2015–2030, 2015. doi: 10.1130/GES01189.1.

Diederichs, A., Nissen, E. K., Lajoie, L. J., Langridge, R. M., Malireddi, S. R., Clark, K. J., Hamling, I. J., and Tagliasacchi, A. Unusual kinematics of the Papatea fault (2016 Kaikōura earthquake) suggest anelastic rupture. Science Advances, 5(10): eaax5703, 2019. doi: 10.1126/sciadv.aax5703.

DiFrancesco, P.-M., Bonneau, D., and Hutchinson, D. J. The implications of M3C2 projection diameter on 3D semi-automated rockfall extraction from sequential terrestrial laser scanning point clouds. Remote Sensing, 12(11):1885, 2020. doi: 10.3390/rs12111885.

DuRoss, C. B., Bunds, M. P., Gold, R. D., Briggs, R. W., Reitman, N. G., Personius, S. F., and Toké, N. A. Variable normal-fault rupture behavior, northern Lost River fault zone, Idaho, USA. Geosphere, 15(6):1869–1892, 2019. doi: 10.1130/GES02096.1.

Eberhart-Phillips, D., Haeussler, P. J., Freymueller, J. T., Frankel, A. D., Rubin, C. M., Craw, P., Ratchkovski, N. A., Anderson, G., Carver, G. A., Crone, A. J., Dawson, T. E., Fletcher, H., Hansen, R., Harp, E. L., Harris, R. A., Hill, D. P., Hreinsdóttir, S., Jibson, R. W., Jones, L. M., Kayen, R., Keefer, D. K., Larsen, C. F., Moran, S. C., Personius, S. F., Plafker, G., Sherrod, B., Sieh, K., Sitar, N., and Wallace, W. K. The 2002 Denali Fault Earthquake, Alaska: A Large Magnitude, Slip-Partitioned Event. Science, 300(5622): 1113–1118, 2003. doi: 10.1126/science.1082703.

Elliott, J., Nissen, E., England, P., Jackson, J. A., Lamb, S., Li, Z., Oehlers, M., and Parsons, B. Slip in the 2010–2011 Canterbury earthquakes, New Zealand. Journal of Geophysical Research: Solid Earth, 117(B3), 2012. doi: 10.1029/2011JB008868.

England, T. and Calon, T. The Cowichan fold and thrust system, Vancouver Island, southwestern British Columbia. Geological Society of America Bulletin, 103(3):336–362, 1991. doi: 10.1130/0016-7606(1991)103<0336:TCFATS>2.3.CO;2.

Esri. Ocean Basemap, 2022. https://www.arcgis.com/home/item.html?id=6348e67824504fc9a62976434bf0d8d5, (December 1, 2023).

European Union Aviation Safety Authority. Drones - National Aviation Authorities, 2022. https://www.easa.europa.eu/en/domains/civil-drones/naa.

Federal Aviation Administration. Become a Drone Pilot, 2023. https://www.faa.gov/uas/commercial_operators/become_a_drone_pilot.

Fernandez-Diaz, J., Carter, W., Shrestha, R., and Glennie, C. Now You See It… Now You Don't: Understanding Airborne Mapping LiDAR Collection and Data Product Generation for Archaeological Research in Mesoamerica. Remote Sensing, 6(10): 9951–10001, 2014. doi: 10.3390/rs6109951.

Finley, T., Salomon, G., Stephen, R., Nissen, E., Cassidy, J., and Menounos, B. Preliminary results and structural interpretations from drone LiDAR surveys over the Eastern Denali fault, Yukon. In Yukon Exploration and Geology, pages 83–105, 2022a. MacFarlane, KE, (ed.), Yukon Geological Survey.

Finley, T. D., Johnston, S. T., Unsworth, M. J., Banks, J., and Pana, D.-I. Modern dextral strain controls active hydrothermal systems in the southeastern Canadian Cordillera. GSA Bulletin, 2022b. doi: 10.1130/B36500.1.

Gabrielse, H., Monger, J., Wheeler, J., and Yorath, C. Morphogeological Belts, Tectonic Assemblages and Terranes, Chapter 2, Part A, of Geology of the Cordilleran Orogen in Canada, Geology of Canada, no. 4. Geological Survey of Canada, pages 15–28, 1991. doi: 10.4095/134073.

GDAL/OGR contributors. GDAL/OGR Geospatial Data Abstraction software Library. Open Source Geospatial Foundation, 2023. doi: 10.5281/zenodo.5884351.

Glennie, C., Brooks, B., Ericksen, T., Hauser, D., Hudnut, K., Foster, J., and Avery, J. Compact Multipurpose Mobile Laser Scanning System — Initial Tests and Results. Remote Sensing, 5(2): 521–538, 2013. doi: 10.3390/rs5020521.

Glennie, C. L., Carter, W. E., Shrestha, R. L., and Dietrich, W. E. Geodetic imaging with airborne LiDAR: the Earth's surface revealed. Reports on Progress in Physics, 76(8):086801, 2013. doi: 10.1088/0034-4885/76/8/086801.

Glennie, C. L., Hinojosa-Corona, A., Nissen, E., Kusari, A., Oskin, M. E., Arrowsmith, J. R., and Borsa, A. Optimization of legacy lidar data sets for measuring near-field earthquake displacements. Geophysical Research Letters, 41(10):3494–3501, 2014. doi: 10.1002/2014GL059919.

Gold, P. O., Oskin, M. E., Elliott, A. J., Hinojosa-Corona, A., Taylor, M. H., Kreylos, O., and Cowgill, E. Coseismic slip variation assessed from terrestrial lidar scans of the El Mayor-Cucapah surface rupture. Earth and Planetary Science Letters, 366:151–162, 2013. doi: 10.1016/j.epsl.2013.01.040.

Gu, Y., Xiao, Z., and Li, X. A Spatial Alignment Method for UAV LiDAR Strip Adjustment in Nonurban Scenes. IEEE Transactions on Geoscience and Remote Sensing, 61:1–13, 2023. doi: 10.1109/TGRS.2023.3281692.

Haddad, D. E., Akciz, S. O., Arrowsmith, J. R., Rhodes, D. D., Oldow, J. S., Zielke, O., Toke, N. A., Haddad, A. G., Mauer, J., and Shilpakar, P. Applications of airborne and terrestrial laser scanning to paleoseismology. Geosphere, 8(4):771–786, 2012. doi: 10.1130/GES00701.1.

Haeussler, P. J., Matmon, A., Schwartz, D. P., and Seitz, G. G. Neotectonics of interior Alaska and the late Quaternary slip rate along the Denali fault system. Geosphere, 13(5):1445–1463, 2017. doi: 10.1130/GES01447.1.

Harrichhausen, N., Morell, K. D., Regalla, C., Bennett, S. E., Leonard, L. J., Lynch, E. M., and Nissen, E. Paleoseismic trenching reveals Late Quaternary kinematics of the Leech River fault: Implications for forearc strain accumulation in northern Cascadia. Bulletin of the Seismological Society of America, 111(2): 1110–1138, 2021. doi: 10.1785/0120200204.

Harrichhausen, N., Morell, K. D., Regalla, C., Lynch, E. M., and Leonard, L. J. Eocene Terrane Accretion in Northern Cascadia Recorded by Brittle Left-Lateral Slip on the San Juan Fault. Tectonics, 41(10):e2022TC007317, 2022. doi: 10.1029/2022TC007317. Publisher: John Wiley & Sons, Ltd.

Harrichhausen, N., Finley, T., Morell, K. D., Regalla, C., Bennett, S.

E. K., Leonard, L. J., Nissen, E., McLeod, E., Lynch, E. M., Salomon, G., and Sethanant, I. Discovery of an Active Forearc Fault in an Urban Region: Holocene Rupture on the XEOLXELEK-Elk Lake Fault, Victoria, British Columbia, Canada. Tectonics, 42 (12):e2023TC008170, 2023. doi: 10.1029/2023TC008170.

Harris Aerial. Carrier H6 Hybrid - Heavy Lift Drones, July 2023. https://harrisaerial.com/carrier-drones/carrier-h6-hybrid/.

Harwin, S. and Lucieer, A. Assessing the Accuracy of Georeferenced Point Clouds Produced via Multi-View Stereopsis from Unmanned Aerial Vehicle (UAV) Imagery. Remote Sensing, 4(6): 1573–1599, 2012. doi: 10.3390/rs4061573.

Haugerud, R. A., Harding, D. J., Johnson, S. Y., Harless, J. L., Weaver, C. S., and Sherrod, B. L. High-resolution lidar topography of the Puget Lowland, Washington. GSA Today, 13(6):4–10, 2003. doi: 10.1130/1052-5173(2003)13<0004:HLTOTP>2.0.CO;2.

Hilley, G. E., DeLong, S., Prentice, C., Blisniuk, K., and Arrowsmith, J. Morphologic dating of fault scarps using airborne laser swath mapping (ALSM) data. Geophysical Research Letters, 37(4): L04301, 2010. doi: 10.1029/2009GL042044.

Hodge, M., Biggs, J., Fagereng, Å., Elliott, A., Mdala, H., and Mphepo, F. A semi-automated algorithm to quantify scarp morphology (SPARTA): application to normal faults in southern Malawi. Solid Earth, 10(1):27–57, 2019. doi: 10.5194/se-10-27-2019.

Hodgson, M. E. and Bresnahan, P. Accuracy of airborne lidar-derived elevation. Photogrammetric Engineering & Remote Sensing, 70(3):331–339, 2004. doi: 10.14358/pers.70.3.331.

Hubbard, T. D., Koehler, R. D., and Combellick, R. A. High-resolution lidar data for Alaska infrastructure corridors. Alaska Division of Geological and Geophysical Surveys: Fairbanks, AK, USA, 3:291, 2011.

Hunter, L. E., Howle, J. F., Rose, R. S., and Bawden, G. W. LiDAR-Assisted Identification of an Active Fault near Truckee, California. Bulletin of the Seismological Society of America, 101(3): 1162–1181, 2011. doi: 10.1785/0120090261.

Isenburg, M. LAStools-efficient LiDAR processing software (version 210418), obtained from https://lastools.github.io/, 2021.

Ishimura, D., Toda, S., Mukoyama, S., Homma, S., Yamaguchi, K., and Takahashi, N. 3D Surface Displacement and Surface Ruptures Associated with the 2014 Mw 6.2 Nagano Earthquake Using Differential Lidar. Bulletin of the Seismological Society of America, 109(2):780–796, 2019. doi: 10.1785/0120180020.

James, M. R. and Robson, S. Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. Journal of Geophysical Research: Earth Surface, 117(F3):F03017, 2012. doi: 10.1029/2011JF002289.

Johnson, K., Nissen, E., Saripalli, S., Arrowsmith, J. R., McGarey, P., Scharer, K., Williams, P., and Blisniuk, K. Rapid mapping of ultrafine fault zone topography with structure from motion. Geosphere, 10(5):969–986, 2014. doi: 10.1130/GES01017.1.

Johnson, K. L., Nissen, E., and Lajoie, L. Surface Rupture Morphology and Vertical Slip Distribution of the 1959 M_w 7.2 Hebgen Lake (Montana) Earthquake From Airborne Lidar Topography. Journal of Geophysical Research (Solid Earth), 123(9): 8229–8248, 2018. doi: 10.1029/2017JB015039.

Johnson, S. Y. Evidence for a margin-truncating transcurrent fault (pre-late Eocene) in western Washington. Geology, 12(9):538–541, 1984. doi: 10.1130/0091-7613(1984)12<538:EFAMTF>2.0.CO;2.

Jones, R. R., Kokkalas, S., and McCaffrey, K. J. W. Quantitative analysis and visualization of nonplanar fault surfaces using terrestrial laser scanning (LIDAR)–The Arkitsa fault, central Greece, as a case study. Geosphere, 5(6):465–482, 2009. doi: 10.1130/GES00216.1.

Kellner, J. R., Armston, J., Birrer, M., Cushman, K. C., Duncanson, L., Eck, C., Falleger, C., Imbach, B., Král, K., Krůček, M., Trochta, J., Vrška, T., and Zgraggen, C. New Opportunities for Forest Remote Sensing Through Ultra-High-Density Drone Lidar. Surveys in Geophysics, 40(4):959–977, 2019. doi: 10.1007/s10712-019-09529-9.

Kolaj, M., Adams, J., and Halchuk, S. The 6th generation seismic hazard model of Canada. Geological Survey of Canada, Open File 8630, pages 1–12, 2020. doi: https://doi.org/10.4095/327322.

Lague, D., Brodu, N., and Leroux, J. Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (N-Z). ISPRS Journal of Photogrammetry and Remote Sensing, 82:10–26, 2013. doi: 10.1016/j.isprsjprs.2013.04.009.

Lajoie, L. J., Nissen, E., Johnson, K. L., Arrowsmith, J. R., Glennie, C. L., Hinojosa-Corona, A., and Oskin, M. E. Extent of Low-Angle Normal Slip in the 2010 El Mayor-Cucapah (Mexico) Earthquake From Differential Lidar. Journal of Geophysical Research (Solid Earth), 124(1):943–956, 2019. doi: 10.1029/2018JB016828.

Langridge, R. M., Ries, W. F., Farrier, T., Barth, N. C., Khajavi, N., and De Pascale, G. P. Developing sub 5-m LiDAR DEMs for forested sections of the Alpine and Hope faults, South Island, New Zealand: Implications for structural interpretations. Journal of Structural Geology, 64:53–66, 2014. doi: 10.1016/j.jsg.2013.11.007.

Leandro, R. F., Santos, M. C., and Langley, R. B. Analyzing GNSS data in precise point positioning software. GPS Solutions, 15 (1):1–13, 2011. doi: 10.1007/s10291-010-0173-9.

Liang, Y., Zhao, C.-Z., Yuan, H., Chen, Y., Zhang, W., Huang, J.-Q., Yu, D., Liu, Y., Titirici, M.-M., Chueh, Y.-L., Yu, H., and Zhang, Q. A review of rechargeable batteries for portable electronic devices. InfoMat, 1(1):6–32, 2019. doi: 10.1002/inf2.12000.

LidarBC. Open LiDAR Data Portal, 2023. https://lidar.gov.bc.ca/.

Lin, Z., Kaneda, H., Mukoyama, S., Asada, N., and Chiba, T. Detection of subtle tectonic–geomorphic features in densely forested mountains by very high-resolution airborne LiDAR survey. Geomorphology, 182:104–115, 2013. doi: 10.1016/j.geomorph.2012.11.001.

Marechal, A., Ritz, J.-F., Ferry, M., Mazzotti, S., Blard, P.-H., Braucher, R., and Saint-Carlier, D. Active tectonics around the Yakutat indentor: New geomorphological constraints on the eastern Denali, Totschunda and Duke River Faults. Earth and Planetary Science Letters, 482:71–80, 2018. doi: 10.1016/j.epsl.2017.10.051.

Meigs, A. Active tectonics and the LiDAR revolution. Lithosphere, 5(2):226–229, 2013. doi: 10.1130/RF.L004.1.

Metcalf, A., Welles, T., Murakami, Y., Nakamura, H., and Ahn, J. Unmanned Aerial Vehicle Solid Oxide Fuel Cell and Internal Combustion Engine Hybrid Powertrain: An Experimental and Simulation Centered Review. American Society of Mechanical Engineers Digital Collection, 2022. doi: 10.1115/POWER2022-86357.

Morell, K. D., Regalla, C., Leonard, L. J., Amos, C., and Levson, V. Quaternary rupture of a crustal fault beneath Victoria, British Columbia, Canada. GSA Today, 7(3), 2017. doi: 10.1130/GSATG291A.1.

Morin, P., Porter, C., Cloutier, M., Howat, I., Noh, M.-J., Willis, M., Bates, B., Willamson, C., and Peterman, K. ArcticDEM; A Publically Available, High Resolution Elevation Model of the Arctic. In EGU General Assembly Conference Abstracts, pages EPSC2016–8396, 2016.

Nash, D. B. Morphologic Dating of Degraded Normal Fault Scarps. Journal of Geology, 88(3):353–360, 1980. doi: 10.1086/628513.

Nelson, A. R., Personius, S. F., Wells, R. E., Schermer, E. R., Bradley, L., Buck, J., and Reitman, N. Holocene Earthquakes of Magnitude 7 during Westward Escape of the Olympic Mountains, Washington. Bulletin of the Seismological Society of America, 107(5):2394–2415, 2017. doi: 10.1785/0120160323.

Nevitt, J. M., Brooks, B. A., Catchings, R. D., Goldman, M. R., Ericksen, T. L., and Glennie, C. L. Mechanics of near-field deformation during co- and post-seismic shallow fault slip. Scientific Reports, 10:5031, 2020. doi: 10.1038/s41598-020-61400-9.

Nissen, E., Maruyama, T., Ramon Arrowsmith, J., Elliott, J. R., Krishnan, A. K., Oskin, M. E., and Saripalli, S. Coseismic fault zone deformation revealed with differential lidar: Examples from Japanese M_w ∼7 intraplate earthquakes. Earth and Planetary Science Letters, 405:244–256, 2014. doi: 10.1016/j.epsl.2014.08.031.

Oskin, M. E., Arrowsmith, J. R., Corona, A. H., Elliott, A. J., Fletcher, J. M., Fielding, E. J., Gold, P. O., Garcia, J. J. G., Hudnut, K. W., Liu-Zeng, J., and Teran, O. J. Near-Field Deformation from the El Mayor-Cucapah Earthquake Revealed by Differential LIDAR. Science, 335(6069):702, 2012. doi: 10.1126/science.1213778.

Pellicani, R., Argentiero, I., Manzari, P., Spilotro, G., Marzo, C., Ermini, R., and Apollonio, C. UAV and Airborne LiDAR Data for Interpreting Kinematic Evolution of Landslide Movements: The Case Study of the Montescaglioso Landslide (Southern Italy). Geosciences, 9(6):248, 2019. doi: 10.3390/geosciences9060248.

Pingel, T. J., Clarke, K. C., and McBride, W. A. An improved simple morphological filter for the terrain classification of airborne LIDAR data. ISPRS Journal of Photogrammetry and Remote Sensing, 77:21–30, 2013. doi: 10.1016/j.isprsjprs.2012.12.002.

Piriz, R., Mozo, A., Navarro, P., and Rodriguez, D. MagicGNSS: Precise GNSS products out of the box. In Proceedings of the 21st International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2008), pages 1242–1251, 2008.

Prentice, C. S., Crosby, C. J., Whitehill, C. S., Arrowsmith, J. R., Furlong, K. P., and Phillips, D. A. Illuminating Northern California's Active Faults. EOS Transactions, 90(7):55, 2009. doi: 10.1029/2009EO070002.

Purba, J. C. S., Gilbert, H., and Dettmer, J. Structure and Dynamics of the Southern Rocky Mountain Trench near Valemount, British Columbia, Inferred from Local Seismicity. Seismological Research Letters, 92(5):3087–3099, 2021. doi: 10.1785/0220200350.

Rajashekara, K. Present Status and Future Trends in Electric Vehicle Propulsion Technologies. IEEE Journal of Emerging and Selected Topics in Power Electronics, 1(1):3–10, 2013. doi: 10.1109/JESTPE.2013.2259614.

Resop, J. P., Lehmann, L., and Hession, W. C. Drone laser scanning for modeling riverscape topography and vegetation: Comparison with traditional aerial lidar. Drones, 3(2):35, 2019. doi: 10.3390/drones3020035.

Risbøl, O. and Gustavsen, L. LiDAR from drones employed for mapping archaeology–Potential, benefits and challenges. Archaeological Prospection, 25(4):329–338, 2018. doi: 10.1002/arp.1712.

Ristau, J., Rogers, G. C., and Cassidy, J. F. Stress in western Canada from regional moment tensor analysis. Canadian Journal of Earth Sciences, 44(2):127–148, 2007. doi: 10.1139/e06-057.

Rusmore, M. E. and Cowan, D. S. Jurassic–Cretaceous rock units along the southern edge of the Wrangellia terrane on Vancouver Island. Canadian Journal of Earth Sciences, 22(8):1223–1232, 1985. doi: 10.1139/e85-124.

Salisbury, J. B., Rockwell, T. K., Middleton, T. J., and Hudnut, K. W. LiDAR and Field Observations of Slip Distribution for the Most Recent Surface Ruptures along the Central San Jacinto Fault. Bulletin of the Seismological Society of America, 102(2): 598–619, 2012. doi: 10.1785/0120110068.

Salomon, G. W., New, T., Muir, R. A., Whitehead, B., Scheiber-Enslin, S., Smit, J., Stevens, V., Kahle, B., Kahle, R., Eckardt, F. D., and Alastair Sloan, R. Geomorphological and geophysical analyses of the Hebron Fault, SW Namibia: implications for stable continental region seismic hazard. Geophysical Journal International, 229(1):235–254, 2022. doi: 10.1093/gji/ggab466.

Sawicki, O. and Smith, D. G. Glacial Lake Invermere, upper Columbia River valley, British Columbia: a paleogeographic reconstruction. Canadian Journal of Earth Sciences, 29(4): 687–692, 1992. doi: 10.1139/e92-059.

Schermer, E. R., Amos, C. B., Duckworth, W. C., Nelson, A. R., Angster, S., Delano, J., and Sherrod, B. L. Postglacial Mw 7.0-7.5 Earthquakes on the North Olympic Fault Zone, Washington. Bulletin of the Seismological Society of America, 111(1): 490–513, 2021. doi: 10.1785/0120200176.

Scott, C., Phan, M., Nandigam, V., Crosby, C., and Arrowsmith, J. R. Measuring change at Earth's surface: On-demand vertical and three-dimensional topographic differencing implemented in OpenTopography. Geosphere, 17(4):1318–1332, 2021. doi: 10.1130/GES02259.1.

Scott, C. P., Arrowsmith, J. R., Nissen, E., Lajoie, L., Maruyama, T., and Chiba, T. The M7 2016 Kumamoto, Japan, Earthquake: 3-D Deformation Along the Fault and Within the Damage Zone Constrained From Differential Lidar Topography. Journal of Geophysical Research (Solid Earth), 123(7):6138–6155, 2018. doi: 10.1029/2018JB015581.

Scott, C. P., DeLong, S. B., and Arrowsmith, J. R. Distribution of Aseismic Deformation Along the Central San Andreas and Calaveras Faults From Differencing Repeat Airborne Lidar. Geophysical Research Letters, 47(22):e90628, 2020. doi: 10.1029/2020GL090628.

Scott, C. P., Beckley, M., Phan, M., Zawacki, E., Crosby, C., Nandigam, V., and Arrowsmith, R. Statewide USGS 3DEP Lidar Topographic Differencing Applied to Indiana, USA. Remote Sensing, 14(4):847, 2022. doi: 10.3390/rs14040847.

Skyfront. Perimeter 8 UAS, 2023. https://skyfront.com/perimeter-8.

Stöcker, C., Bennett, R., Nex, F., Gerke, M., and Zevenbergen, J. Review of the Current State of UAV Regulations. Remote Sensing, 9(5), 2017. doi: 10.3390/rs9050459.

Telling, J., Lyda, A., Hartzell, P., and Glennie, C. Review of Earth science research using terrestrial laser scanning. Earth Science Reviews, 169:35–68, 2017. doi: 10.1016/j.earscirev.2017.04.007.

Tomsett, C. and Leyland, J. Development and Testing of a UAV Laser Scanner and Multispectral Camera System for Eco-Geomorphic Applications. Sensors, 21(22):7719, 2021. doi: 10.3390/s21227719.

Toth, C., Brzezinska, D., Csanyi, N., Paska, E., and Yastikli, N. LiDAR mapping supporting earthquake research of the San Andreas fault. In Proceedings of the ASPRS 2007 Annual Conference, pages 1–11. American Society for Photogrammetry and Remote Sensing (ASPRS) Tampa, FL, 2007.

Townsend, A., Jiya, I. N., Martinson, C., Bessarabov, D., and Gouws, R. A comprehensive review of energy sources for unmanned aerial vehicles, their shortfalls and opportunities for improvements. Heliyon, 6(11):e05285, 2020. doi: 10.1016/j.heliyon.2020.e05285.

Transport Canada. Getting a drone pilot certificate, 2022. https://tc.canada.ca/en/aviation/drone-safety/drone-pilot-licensing/getting-drone-pilot-certificate.

Transport Canada. Minister of Transport announces Canada's first proposed drone safety regulations for beyond visual line-of-sight operations, 2023. https://www.canada.ca/en/transport-canada/news/2023/06/minister-of-transport-announces-canadas-first-proposed-drone-safety-regulations-for-beyond-visual-line-of-sight-operations.html.

UK Civil Aviation Authority. Registering a drone or model aircraft, 2023. https://register-drones.caa.co.uk/individual.

van der Velden, A. J. and Cook, F. A. Structure and tectonic development of the southern Rocky Mountain trench. Tectonics, 15 (3):517–544, 1996. doi: 10.1029/95TC03288.

Van Tassel, C. Defining the true cost behind implementing lidar systems into your business., 2021. https://candrone.com/blogs/news/the-real-cost-of-starting-a-lidar-drone-business.

VanValkenburgh, P., Cushman, K., Butters, L. J. C., Vega, C. R., Roberts, C. B., Kepler, C., and Kellner, J. Lasers without lost cities: Using drone lidar to capture architectural complexity at Kuelap, Amazonas, Peru. Journal of Field Archaeology, 45 (sup1):S75–S88, 2020. doi: 10.1080/00934690.2020.1713287.

Viswanathan, V., Epstein, A. H., Chiang, Y.-M., Takeuchi, E., Bradley, M., Langford, J., and Winter, M. The challenges and opportunities of battery-powered flight. Nature, 601(7894):519–525, 2022. doi: 10.1038/s41586-021-04139-1.

Wang, S., Ren, Z., Wu, C., Lei, Q., Gong, W., Ou, Q., Zhang, H., Ren, G., and Li, C. DEM generation from Worldview-2 stereo imagery and vertical accuracy assessment for its application in active tectonics. Geomorphology, 336:107–118, 2019. doi: 10.1016/j.geomorph.2019.03.016.

Wedmore, L. N. J., Gregory, L. C., McCaffrey, K. J. W., Goodall, H., and Walters, R. J. Partitioned Off-Fault Deformation in the 2016 Norcia Earthquake Captured by Differential Terrestrial Laser Scanning. Geophysical Research Letters, 46(6): 3199–3205, 2019. doi: 10.1029/2018GL080858.

Wei, Z., He, H., Su, P., Zhuang, Q., and Sun, W. Investigating paleoseismicity using fault scarp morphology of the Dushanzi Reverse Fault in the northern Tian Shan, China. Geomorphology, 327:542–553, 2019. doi: 10.1016/j.geomorph.2018.11.025.

Westoby, M., Brasington, J., Glasser, N., Hambrey, M., and Reynolds, J. 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology, 179:300–314, 2012. doi: 10.1016/j.geomorph.2012.08.021.

Wiatr, T., Reicherter, K., Papanikolaou, I., Fernández-Steeger, T., and Mason, J. Slip vector analysis with high resolution t-LiDAR scanning. Tectonophysics, 608:947–957, 2013. doi: 10.1016/j.tecto.2013.07.024.

Wieser, M., Hollaus, M., Mandlburger, G., Glira, P., and Pfeifer, N. ULS LiDAR supported analyses of laser beam penetration from different ALS systems into vegetation. ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences, 3(3), 2016. doi: 10.5194/isprs-annals-III-3-233-2016.

Witter, R. C., Bender, A. M., Scharer, K. M., DuRoss, C. B., Haeussler, P. J., and Lease, R. O. Geomorphic expression and slip rate of the Fairweather fault, southeast Alaska, and evidence for predecessors of the 1958 rupture. Geosphere, 17(3):711–738, 2021. doi: 10.1130/GES02299.1.

Xiaoye Liu. Airborne LiDAR for DEM generation: some critical issues. Progress in Physical Geography: Earth and Environment, 32(1):31–49, 2008. doi: 10.1177/0309133308089496.

Zhang, B., Liao, Y., Guo, S., Wallace, R. E., Bucknam, R. C., and Hanks, T. C. Fault scarps related to the 1739 earthquake and seismicity of the Yinchuan graben, Ningxia Huizu Zizhiqu, China. Bulletin of the Seismological Society of America, 76(5): 1253–1287, 1986. doi: 10.1785/BSSA0760051253.

Zhang, K., Chen, S.-C., Whitman, D., Shyu, M.-L., Yan, J., and Zhang, C. A progressive morphological filter for removing nonground measurements from airborne LIDAR data. IEEE Transactions on Geoscience and Remote Sensing, 41(4):872–882, 2003. doi: 10.1109/TGRS.2003.810682.

Zhang, W., Qi, J., Wan, P., Wang, H., Xie, D., Wang, X., and Yan, G. An Easy-to-Use Airborne LiDAR Data Filtering Method Based on Cloth Simulation. Remote Sensing, 8(6), 2016. doi: 10.3390/rs8060501.

Zhu, X., Glennie, C. L., and Brooks, B. A. Automated near-field deformation detection from mobile laser scanning for the 2014 M_w 6.0 South Napa earthquake. Journal of Applied Geodesy, 16(1): 65–79, 2022. doi: 10.1515/jag-2021-0023.

Zielke, O., Arrowsmith, J. R., Ludwig, L. G., and Akçiz, S. O. Slip in the 1857 and Earlier Large Earthquakes Along the Carrizo Plain, San Andreas Fault. Science, 327(5969):1119–1122, 2010. doi: 10.1126/science.1182781.

Zielke, O., Klinger, Y., and Arrowsmith, J. R. Fault slip and earthquake recurrence along strike-slip faults - Contributions of high-resolution geomorphic data. Tectonophysics, 638:43–62, 2015. doi: 10.1016/j.tecto.2014.11.004.

doi:10.26443/seismica.v3i1.1185

SEISMICA

Deep learning detects uncataloged low-frequency earthquakes across regions

Jannes Münchmeyer * [1], **Sophie Giffard-Roisin** [1], **Marielle Malfante** [2], **William B. Frank** [3], **Piero Poli** [4], **David Marsan**[1], **Anne Socquet** [1]

[1]Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, Univ. Gustave Eiffel, ISTerre, Grenoble, France, [2]Univ. Grenoble Alpes, CEA, List, Grenoble, France, [3]Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA, [4]Dipartimento di Geoscienze, Università di Padova, Padova, Italy

Abstract Documenting the interplay between slow deformation and seismic ruptures is essential to understand the physics of earthquakes nucleation. However, slow deformation is often difficult to detect and characterize. The most pervasive seismic markers of slow slip are low-frequency earthquakes (LFEs) that allow resolving deformation at minute-scale. Detecting LFEs is hard, due to their emergent onsets and low signal-to-noise ratios, usually requiring region-specific template matching approaches. These approaches suffer from low flexibility and might miss LFEs as they are constrained to sources identified a priori. Here, we develop a deep learning-based workflow for LFE detection and location, modeled after classical earthquake detection with phase picking, phase association, and location. Across three regions with known LFE activity, we detect LFEs from both previously cataloged sources and newly identified sources. Furthermore, the approach is transferable across regions, enabling systematic studies of LFEs in regions without known LFE activity.

Non-technical summary Earthquakes are caused by sudden movements on tectonic faults. While such sudden movements have been documented for thousands of years, the last decades have revealed that tectonic faults also host a wide range of slow deformation. Such slow slip happens over the scale of days to years but is still substantially faster than regular plate convergence rates. Recent years have shown that slow slip can play an essential role in the buildup of large earthquakes. Classically, slow deformation is detected and characterised using geodetic observations, such as GNSS or InSAR. This limits the time and space resolution. An alternative is looking for seismic markers accompanying slow slip, among which the most pervasive are low-frequency earthquakes (LFE). Due to their low signal to noise ratio and emergent onsets, such LFEs are notoriously difficult to detect. Here, we develop a novel method for detecting LFEs using deep learning. Our method successfully detects LFEs from both known and unknown sources. In contrast to previous approaches, our method can detect LFEs without prior knowledge of the region, which makes it promising for LFE detection in regions where no LFEs have been found previously.

Production Editor:
Gareth Funning
Handling Editor:
Jack Muir
Copy & Layout Editor:
Abhineet Gupta

Signed reviewer(s):
David Shelly

Received:
January 18, 2024
Accepted:
April 10, 2024
Published:
May 10, 2024

1 Introduction

Stress release on tectonic faults can happen in two ways: fast and slow. Fast deformation happens in the form of earthquakes; slow relaxation is observed as creep or episodes of accelerated slip, so-called slow slip events (Dragert et al., 2001; Ozawa et al., 2002; Lowry et al., 2001; Ide et al., 2007a). The complex interactions between fast and slow deformation might be at play during the initiation of large earthquakes (Radiguet et al., 2016; Socquet et al., 2017; Cruz-Atienza et al., 2021). However, studying these interactions requires detailed catalogs of both deformation types. While the impulsive nature of earthquakes causes clear signatures on seismic recordings, detecting slow slip is substantially more challenging. Its detection commonly uses geodetic observations with a limited spatial and temporal resolution (Michel et al., 2019; Okada et al., 2022; Costantino et al., 2023).

An alternative way to map slow deformation is by detecting and characterising its seismic markers. One such type of markers are low-frequency earthquakes (LFEs), weak seismic signals with a duration on the scale of seconds. Recent research shows that the rate and magnitude of LFEs track the slow deformation (Frank and Brodsky, 2019; Mouchon et al., 2023). LFEs are similar to regular earthquakes in some characteristics, e.g., distinct phase arrivals or predominantly double-couple sources, but also have clear differences (Shelly et al., 2007; Ide et al., 2007b; Royer and Bostock, 2014; Imanishi et al., 2016; Supino et al., 2020; Wang et al., 2023). First, they have an eponymous depletion of energy in the high-frequency band (above a few Hz). Second, in consequence of missing high frequencies, they do not exhibit impulsive arrivals but are emergent, making them hard to detect. Third, they often occur in intense bursts with inter-event times of only seconds, leading to superimposed waveforms commonly referred to as tremors (Shelly et al., 2007).

To illustrate the challenges these characteristics cause for LFE detection, it is worth contrasting LFE detection with the identification of regular earthquakes.

*Corresponding author: munchmej@univ-grenoble-alpes.fr

Detecting regular earthquakes traditionally relies on a two-step procedure: (i) phase picking, i.e., identifying P and S waves arrival times at seismic stations; (ii) phase association, i.e., selecting sets of picks across stations that are consistent with a common source location and origin time. Downstream analysis can then determine the event location and additional source parameters. In this workflow, a side benefit of the phase association step is that it acts as a quality control to remove spurious phase picks. At the moment, such a workflow is usually not applicable to LFE detection, as the low signal-to-noise (SNR) ratio and the emergent onsets make it impossible for classical algorithms to pick phase arrivals. There are exceptions, notably the JMA catalog (Japan Meteorological Agency, 2023), but these rely on high-quality and high-density data, manual intervention, and high SNR LFEs. Instead, LFE detection usually relies on manual identification (Shelly, 2010), beamforming (Frank and Shapiro, 2014), or phase coherence (Gombert and Hawthorne, 2023). These approaches often suffer from high computational demand or requirement for manual labour. However, they can be used to generate LFE template waveforms forming an initial catalog for a subsequent matched filtering search on long running recordings. As the initial approaches often fail to identify all existing LFE sources, such catalogs will be biased towards certain sources.

Due to its high sensitivity, matched filtering, also known as template matching, has become the de facto standard for LFE detection (Shelly, 2017; Bostock et al., 2015; Frank et al., 2014). Once initial templates are identified, the method identifies repeat occurrences of the template events by correlating these with the continuous waveforms. In addition to detecting occurrences, this procedure groups the LFEs into families according to their matching templates. This allows to stack waveforms and accurately locate the families. While highly sensitive, matched-filtering presents several disadvantages: templates are always region and station specific, matched filtering does not provide locations for individual events, and the model can not detect LFEs outside the initially detected families. Especially the last limitation shapes our understanding of LFEs, as template matching can only recover repeating events, potentially skewing our view of overall LFE behavior by the most repetitive sources. Furthermore, the grouping into families is partially artificial, as template matches often overlap, i.e., many detections can not be uniquely assigned to one family.

A closely linked task to the detection of LFEs is the detection of tectonic tremors. Typical methods leverage either coherency across station, through source scanning (Kao et al., 2005), waveform coherency (Armbruster et al., 2014), envelope correlations (Bombardier et al., 2023), or repetitiveness of waveform motives within or across stations (Rubin and Armbruster, 2013). However, while the underlying processes are closely related, the tasks pose distinct challenges. Tremors are usually several minutes long, making them easier to detect than LFEs. In addition, these longer waveforms make it easier to locate them, as more characteristics, e.g., envelopes, can be used for location. In contrast, LFEs are short signals, with waveforms lasting at most a few seconds, making detection and location more difficult. However, the short duration of LFEs also implies that they can monitor underlying processes at a higher resolution than tremors, thereby providing additional insights into slow deformation. In some cases, waveform coherency methods similar to tremor detection can be applied for LFE detection, but the results are usually restricted to high signal-to-noise ratio examples (Savard and Bostock, 2015).

To build a flexible LFE detector addressing the disadvantages of template matching, it would be appealing to make a more traditional earthquake detection workflow applicable to LFEs. The critical point for this is a viable automatic phase picker for LFE arrivals. We borrow from the recent breakthroughs in seismic phase picking with deep learning, where recent neural network models have substantially improved earthquake detection (Zhu and Beroza, 2019; Ross et al., 2018; Münchmeyer et al., 2022). These neural network models are trained on millions of manually labelled phase arrivals and thereby learn to accurately discern seismic phase arrivals from noise and accurately determine arrival times. The application of these models to continuous data has allowed to substantially increase the completeness of earthquake catalogs (Tan et al., 2021; González-Vidal et al., 2023; Moutote et al., 2023).

For tremor and LFE detection with deep learning, only few studies exist. Rouet-Leduc et al. (2020) identify tremor episodes in single-station records, but do not attempt to detect or locate individual LFEs. Thomas et al. (2021) focus on LFEs on the San Andreas fault and test model configurations on cataloged events. The preprint of Lin et al. (2023) presents an LFE detection workflow similar to the one presented here but focus exclusively on Southern Vancouver Island. Here, we develop a deep learning based LFE picker and show its applicability to three independent study regions: Cascadia, Guerrero and Nankai. To train our picker, we develop a novel strategy for synthetic data generation that allows for fine-grained control of the training process. Using this method, we set up a classical earthquake detection workflow and demonstrate how to automatically create LFE catalogs across different world regions. Our model successfully identifies and locates individual LFEs, even without using any training examples from the target region. The resulting catalogs are coherent with classical catalogs but have been obtained in a fully automated and region-agnostic manner. Furthermore, the catalogs identify LFEs missing from the reference catalogs, showing that our approach can uncover sources missed in the template matching procedures. We make the trained picker available with a user-friendly interface through SeisBench (Woollam et al., 2022).

2 Training and validation of a deep learning LFE phase picker

For detecting LFEs and determining their phase arrival times, we build a deep learning network. Our network is closely modeled after PhaseNet (Zhu and

Beroza, 2019) due to the model's simplistic architecture and the excellent performance on earthquake data (Münchmeyer et al., 2022). PhaseNet is a 1D U-Net, i.e, a neural network consisting of convolutional encoder and decoder branches and skip connections (Ronneberger et al., 2015). We provide the model with 60 s of 3-component waveforms sampled at 20 Hz, bandpass-filtered between 1 and 8 Hz, the band in which LFEs are typically observed. The model outputs probability curves for P and S phase arrivals. We provide full details on the model and training procedure in the supplement.

In contrast to traditional earthquake pickers, training the model on cataloged LFE waveforms is suboptimal. First, LFEs occur in bursts, i.e., around one LFE arrival there are often further arrivals many of which have not been labelled. This leads to incorrect labelling and in addition makes a quantitative analysis of the model performance difficult. Second, most LFE catalogs are based on template matching, i.e., individual arrivals need to be inferred from arrival times on templates. Due to the low SNR, these times are often highly inaccurate, leading to high model uncertainties. Instead, we train our model on synthetics. For this, we combine LFE stacks with real seismic noise, allowing us to control the number and timing of LFEs and the SNR (Figure 1). We use up to three LFEs per trace to train the model to recognise events with low inter-event times. We use seismic noise from the INSTANCE dataset for Italy (Michelini et al., 2021) as it contains no known LFEs.

We train our model using four regions: Southern Vancouver Island in Cascadia (Canada/USA) (Bostock et al., 2015), the central section of the San Andreas fault (USA) (Shelly, 2017), Guerrero (Mexico) (Frank et al., 2014), and Nankai (Japan) (Japan Meteorological Agency, 2023). Figure S1 shows the distribution of events and stations in the reference catalogs. For Cascadia, San Andreas and Guerrero, we use template matching catalogs and the previously described strategy for generating examples. For Nankai, we apply the classical training scheme as used for earthquakes as individual picks are available. Further details on the datasets can be found in the supplement.

We evaluate our trained models quantitatively on synthetic examples generated with the previously described noise plus stack strategy. The performance on synthetic data can serve as a proxy for the expected performance on real data. We exclude the Nankai catalog from the analysis, because the catalog incompleteness precludes the extraction of challenging yet guaranteed LFE-free time windows. As this study focuses on the generated LFE catalogs, we only provide a synopsis of the analysis on synthetics here and refer to the supplement for further details.

Overall, the models show excellent detection performance for both P and S waves, with an area under the curve (AUC) of the receiver operator characteristics of 0.97 to 1.00 in all regions for positive SNR in dB scale (Figure 1). The performance degrades mildly at -2.5 dB SNR and more sharply after, but all AUC values stay above 0.88 even at -10 dB SNR. Models transfer well across regions with the worst results for a model trained exclusively on Cascadia (Figure S2). The best performing model is the one trained jointly on all four regions. Therefore, we use the model trained jointly on all regions in the subsequent analysis unless explicitly stated otherwise.

Analysing the pick time residuals, clear regional differences are visible, with lowest residuals in Cascadia (Figure S3). In all regions, average residuals are about 0.3 s larger for P arrivals than S arrivals, indicating that these are more difficult to pick. With standard deviations between 0.3 and 1.3 s (at 0 dB), residuals are substantially higher than for traditional earthquake pickers (Münchmeyer et al., 2022). Nonetheless, the residuals expose only low or no bias across all regions. For the regional differences in performance, we think that they can primarily be attributed to the heterogeneity in data quality and SNR. For example, the Cascadia stacks show the highest SNR, leading to the lowest pick residuals. In turn, this implies that no conclusions about inherent regional differences in difficulty for picking LFEs can be inferred.

3 Building deep learning LFE catalogs

Using our phase picking model, we set up an LFE catalog workflow similar to the classical earthquake detection workflow. Here we provide an overview of the workflow with further details in the supplement. First, we pick P and S phases by applying the trained deep learning model to continuous waveforms using SeisBench (Woollam et al., 2022). Second, we use the PyOcto phase associator (Münchmeyer, 2024) to identify coherent arrivals across stations. Third, we use NonLinLoc (Lomax et al., 2000) with a 1D velocity model to perform absolute location of the events. To avoid false detections, we filter the events based on the number of phase picks and the location residuals. For comparison, we create earthquake catalogs using the same waveforms and workflow but using a PhaseNet model trained on INSTANCE implemented in SeisBench as the picker (Zhu and Beroza, 2019; Michelini et al., 2021; Woollam et al., 2022).

As we observed a certain number of events detected as both earthquakes (EQs) and LFEs, we remove these events from the LFE catalogs (Figure S4). We note that it is not clear whether these events should be classified as LFEs or EQs. The level of overlap is dependent on the region, with almost no overlap in Cascadia and Guerrero, a 5% overlap on the San Andreas fault, and a 40% overlap in Nankai. While we are not certain what causes this different behavior, we speculate that in Nankai, LFEs and earthquakes show a wide range of apparent spectra, due to the diverse event distribution and frequency dependent attenuation. This might lead to a higher number of overlapping detections.

We apply our workflow to compile LFE catalogs for the four study regions. As we focus on studying the performance of the model and its resulting catalogs, we restrict ourselves to short study periods: 2003-02-26 to 2003-03-10, 2004-07-02 to 2004-07-27, and 2005-09-03 to 2005-09-25 for Cascadia; 2005-09-01 to 2005-11-30 for Guerrero; 2014-07-01 to 2014-10-01 for San Andreas; 2012-05-25 to 2012-06-14 for Nankai. We chose the pe-

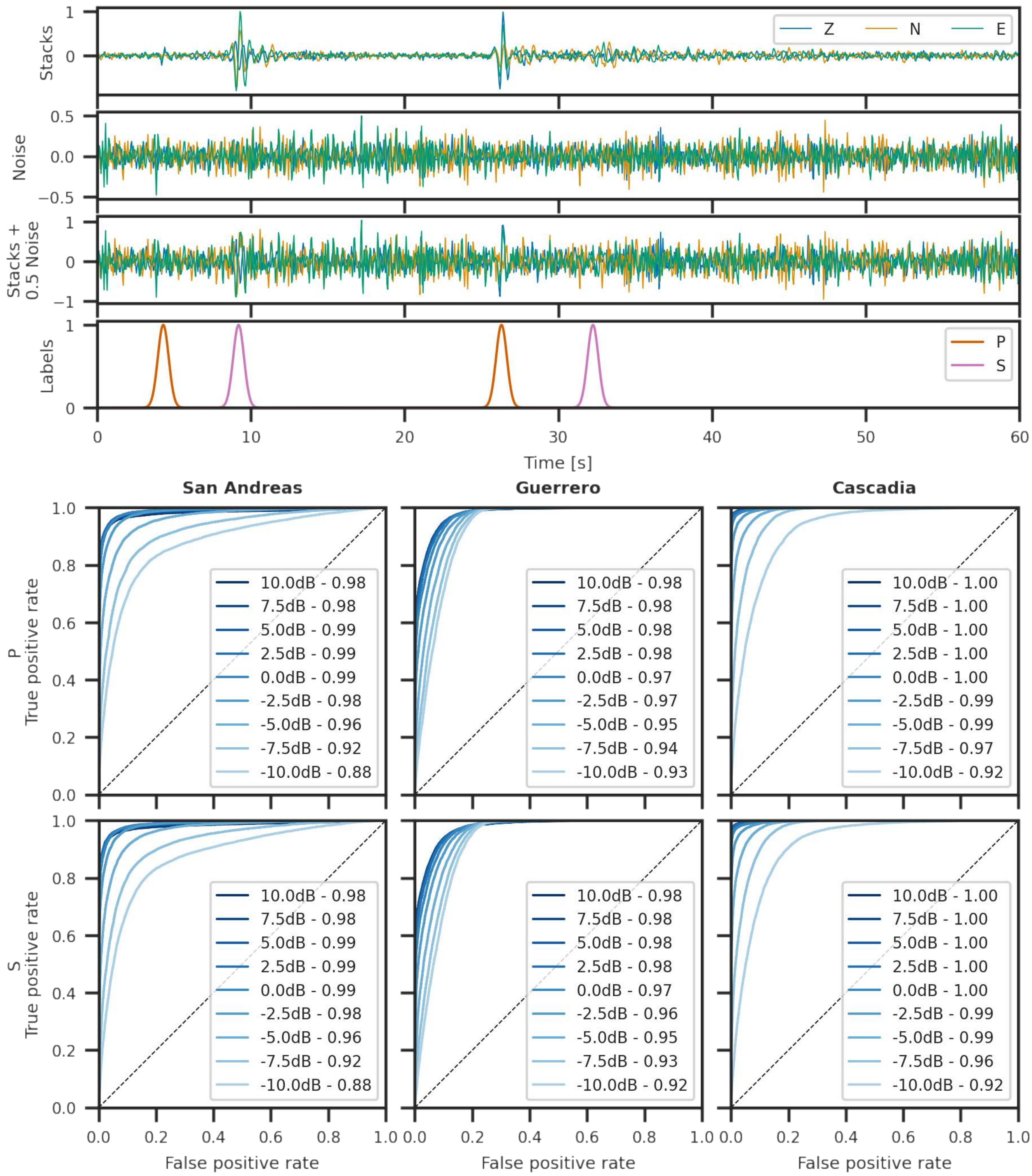


Figure 1 Data generation procedure and evaluation results for synthetic data. The top panels show (top to bottom): the combination of two LFE stacks from Cascadia, a 60 s noise segment from INSTANCE, the combination of signal and noise at 3 dB SNR, and the Gaussian pulse labels for the P and S arrivals. The bottom panels show the receiver operating characteristics (ROC) at different SNR. The numbers in the legend indicate the area under the ROC curve (AUC). For all plots, we use the joint model trained on all four datasets.

riods to contain both intense LFE activity and segments without any identified LFEs.

Figure 2 shows the spatial event distributions. While the overall event locations are scattered, we notice strong similarities with the reference catalogs. For Cascadia (10211 events detected), LFE activity is distributed along a band underneath South Vancouver island. For Guerrero (876 events detected), LFEs occur mostly in a band between 100° and 99° West and around 18.25° North. For Nankai (2525 events detected), a clear band of LFEs is visible in Southwestern Nankai. Further LFEs around 135.5° E match the second band of LFEs com-

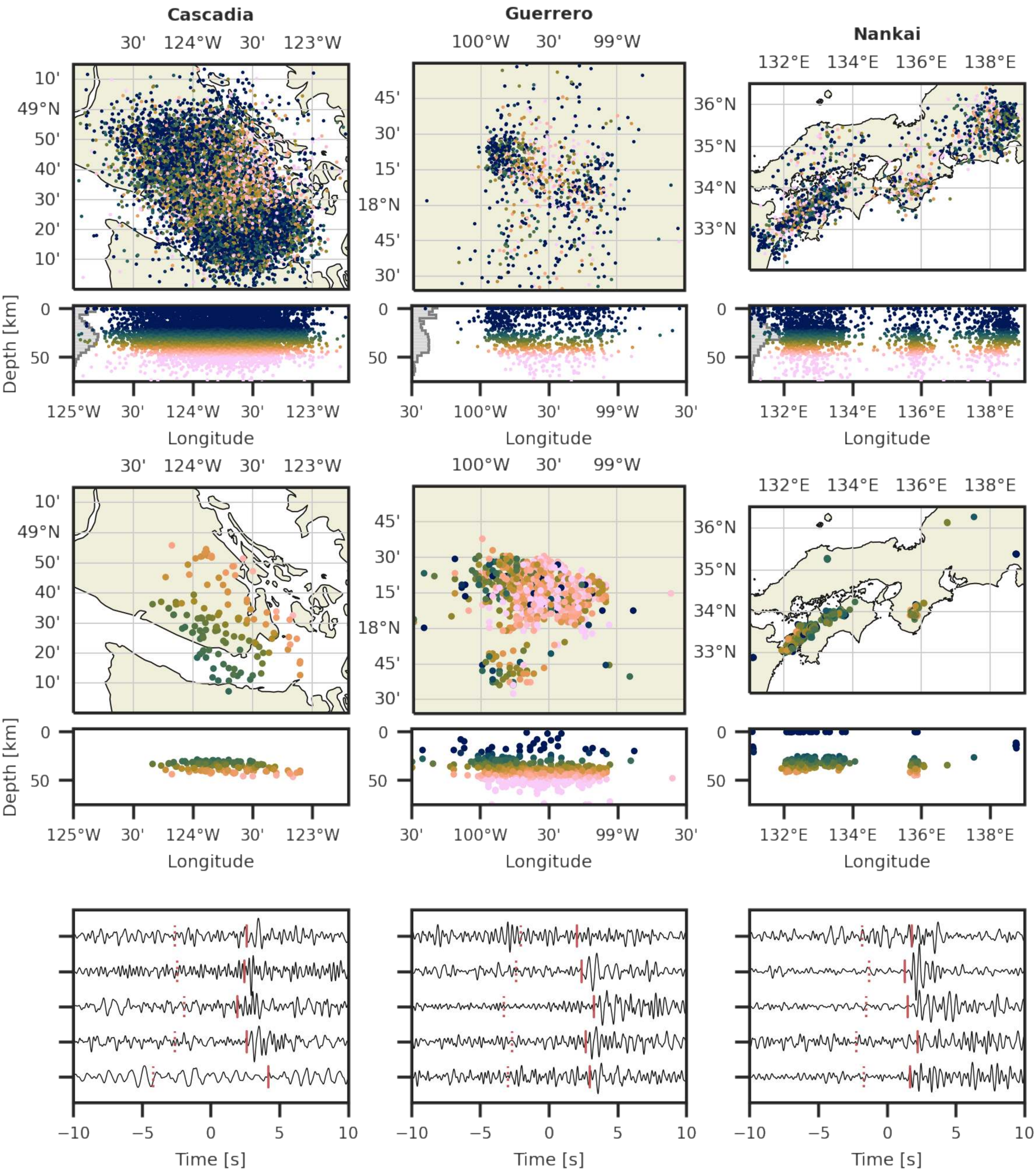


Figure 2 LFE catalogs obtained from deep learning (top row) and the reference catalogs (middle row). For deep learning, each dot represents one LFE. The bottom subpanels show depth cross-sections, showing longitude and depth of events. Color encodes event depth. The histograms on the left of the cross-section show the depth distribution of the detected events. In the reference catalogs for Cascadia (Bostock et al., 2015) and Guerrero (Frank et al., 2014) each dot represents an LFE family. For Nankai (Japan Meteorological Agency, 2023) individual LFEs are plotted. The bottom panels show waveforms of associated LFE picks from deep learning in each region. Red lines indicate phase picks (dotted for P, solid for S).

monly observed in Nankai. For San Andreas (975 events detected), the new catalog deviates from the previous observations, with the detection broadly distributed in space instead of along the fault (Figure S7). This is likely caused by very poor locations due to the station geometry. As many events are only detected by the Parkfield borehole array with very dense station spacing, the aperture is small. Together with high pick uncertainties, this makes determining accurate locations challenging. Therefore, we will exclude San Andreas from the following analysis.

In all catalogs, the event depth exposes high scatter. Nonetheless, the largest density of events is around the previous estimates of LFE source depths. For Cascadia, events within the network show less scatter on depth than outside the network. We suggest that this is caused by the high timing uncertainty of the picks. In particular, a high P pick uncertainty will cause poor depth constraints as the P to S time is indicative of depth. Template matching catalogs alleviate this problem by locating LFE families instead of single events. In Guerrero, we observe an arc-shaped depth distribution. This is most likely related to the station distribution that traces out almost a straight line, leading to poor location constraints perpendicular to the station line (Frank and Shapiro, 2014).

Figure S5 shows the event density of the detected events. This visualisation further highlights the match with the reference catalogs both in terms of latitude and longitude and in terms of depth. For Japan, the highest event density occurs in the South-Western band of LFE activity. For Cascadia, the fine structure of detected LFEs is compatible with earlier publications, e.g., clear overlap is visible with patches A and C in the visualisation of LFE density of Figure 7 by Savard and Bostock (2015).

Even though the overall shape of the catalogs is consistent with the previous catalogs, this alone does not confirm that the identified events are indeed LFEs. We therefore conduct additional analysis into the newly obtained catalogs. Figure 3 shows spatial and temporal patterns in the catalogs. In Cascadia, LFEs in all three observed sequences show a clear North-Westward migration. This is consistent with the slow slip and tremor migration patterns in the area during these episodes (Wech, 2021). We conducted an additional analysis, comparing the PNSN tremor catalog and LFEs detected using our method for a 31 day period in 2021 (Figure S6). This analysis shows a high agreement between LFE and tremor locations, density, and temporal development. This holds even though for this time period we used a less dense station coverage and considered a larger part of Vancouver Island than in the 2003 to 2005 episodes.

For Nankai, we observe a migration in the North-Eastward direction. Notably, the migration is not continuous but rather has a gap and an additional, earlier cluster at the far North-West. This pattern matches exactly the migration pattern in the JMA catalog. We do not observe clear spatial migration patterns in Guerrero, however, such patterns have previously only been identified with very precise location estimates (Frank et al., 2014). In all regions, the evolution of daily event rate between the deep learning and reference catalogs is highly similar with Pearson correlation coefficients between 0.74 and 0.93. In absolute numbers, the deep learning method detects substantially fewer events than the template matching, but more events than the manual detection procedure of the JMA. We note that the number of events from deep learning is highly dependent on the chosen quality control parameters, which we set rather conservatively to avoid false positive detections.

Figure 4 shows a comparison of the velocity spectra of LFEs detected by our method, earthquakes and noise in the three regions. The spectra clearly show the characteristics of the different event types. The earthquake spectra show increasing or at least constant energy up to about 10 Hz. In contrast, the LFEs show a continuous decrease or at most constant levels of energy from low frequencies onward. The LFEs only show substantially higher energy than the noise in a small frequency band, while the EQs show substantially higher SNR at high frequencies. This depletion in energy at higher frequencies is the key property of LFEs.

4 Increased diversity of LFE sources through deep learning

We compare the detected events to the reference catalogs (Figure S8). In Cascadia, for 64% of the LFEs from our workflow, we find an LFE in the reference catalog within 10 s; for Guerrero for 81% of the events. For Nankai, only 8% of our LFEs are in the reference catalog, however our catalog also substantially exceeds the original catalog in the total number of events. Conversely, we recover 39% of the cataloged events. Note that a loose threshold for matching is justified due to uncertainties in the origin times due to inaccurate locations. On one hand, these results are another confirmation that the method correctly identifies LFEs. On the other hand, the substantial fraction of uncataloged events suggests that our method identifies previously unidentified LFEs. In the following, we verify and analyse these detections.

First, we rule out spurious detection. To this end, we scramble the picked arrival times of each station by applying small random shifts. We choose constant shifts for each station for every hour. This destroys the exact times, while keeping the pick distribution, the P to S times within a station, and the higher number of picks within tremor bursts intact. We then build "catalogs" by associating these scrambled pick times, using the same associator settings as for the actual catalogs. The scrambled "catalogs" only contain about 5% to 10% of the number of events contained in the original catalogs and show no spatial coherence (Figure S10). Even these numbers are still likely an overestimation of the false positive rate, as events recorded at many stations are likely to be unperturbed by our scrambling procedure. Therefore, at least 20% additional detections can not be attributed to spurious associations.

Mapping the events without matches in the reference catalog (Figure S11) reveals that they follow the same spatial extent and migration pattern as the full catalog. Notably, for Cascadia and Mexico there are changes in the temporal patterns. For Cascadia, the newly detected events concentrate early and late in the LFE sequence, coinciding with a spatial location around the southeastern tip of Vancouver island and towards the northwestern end of the LFE cluster. Nonetheless, there are additional detections dispersed throughout the whole region including the central region with good coverage in the reference catalog. For Mexico, the largest frac-

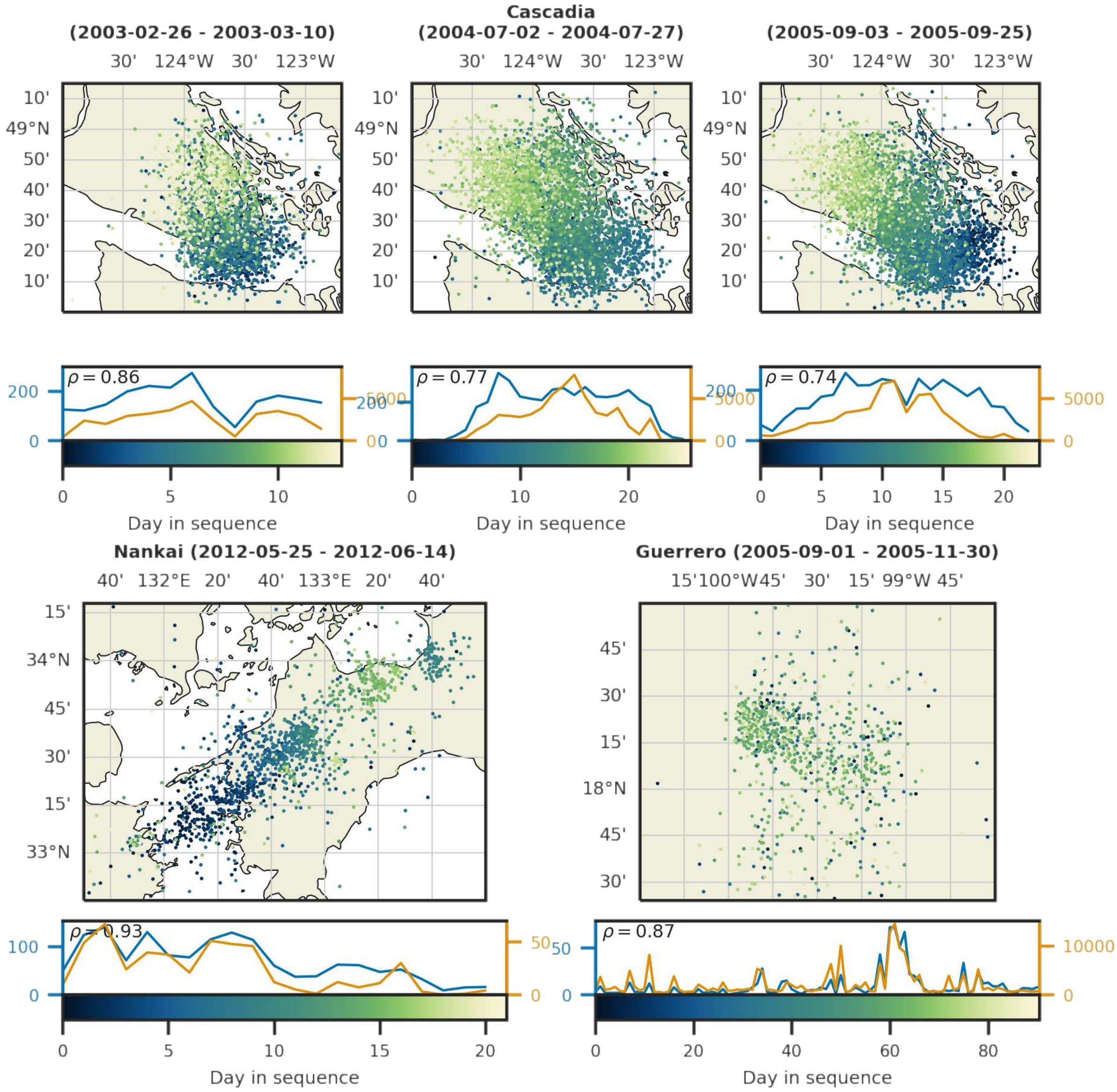


Figure 3 Spatial and temporal migration patterns of detected LFEs. Each dot represents one LFE. The time within the sequence is indicated by colour. The bottom panel shows the number of events per day for the LFEs from deep learning (blue) and reference catalogs (orange). The numbers in the upper left corners indicate the Pearson correlation coefficient between the daily number of events between the two catalogs.

tion of new detections clusters in time between days 30 and 50 of the analyses sequence. Visualising the interevent time (Figure 4) confirms these observations. Both the full deep learning catalogs and the catalog of events without a match in the reference catalogs show clear burst behaviour. In particular for Mexico, certain LFE bursts are contained virtually completely in the template matching catalog, while others have not been identified at all. This highlights that the newly detected LFEs do not only uncover new sources but even new LFE bursts.

To further validate this finding, we correlate the uncataloged detections with the family stacks from the reference catalogs (Figure S12). For Cascadia, the distribution of correlation values for these uncataloged detections are identical to the noise distribution, i.e., the new detections do not match known sources. For Guerrero, some events show systematically higher correlations than the noise. Nonetheless, the vast majority of our additional detections do not match the known sources any better than the noise. This verifies that these new detections are systematically different and that these events can not be found with template matching without identifying further, novel templates.

Extending upon the finding that the model can generalise from known families to LFEs outside these families, we investigate the ability to detect LFEs in regions the model has not been trained on. For this analysis, we trained leave-one-out models, i.e., models trained on all but one region, and applied them on the left-out region. Figure S13 visualises the spatial and temporal migration patterns. Again, the clear migration patterns in Nankai and Cascadia are retrieved. Furthermore, the number of events correlates highly (Pearson correlation between 0.69 and 0.82) with the reference catalogs. The total size of the catalogs varies, with a substantially

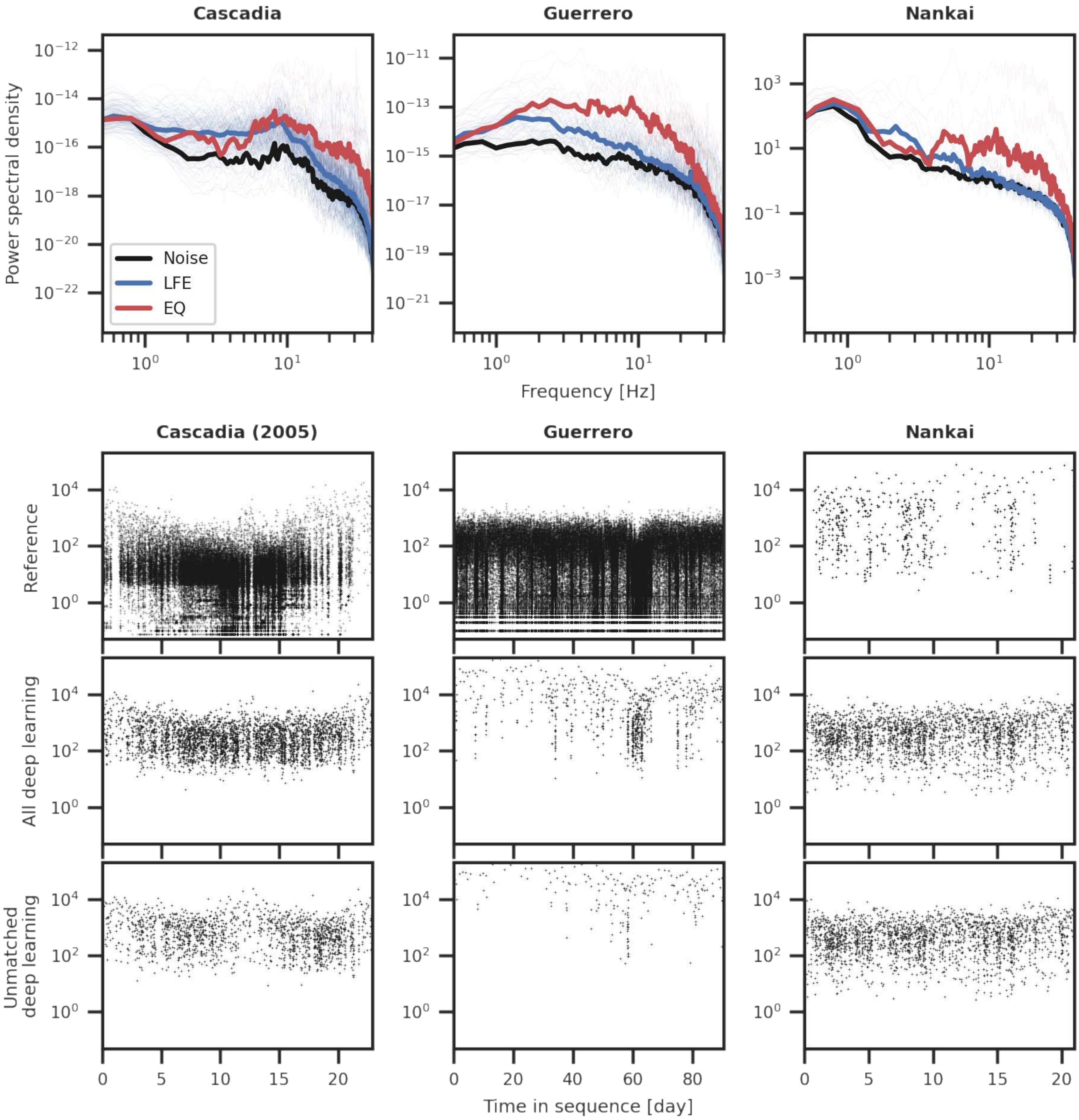


Figure 4 Spectra and interevent times in the different regions. The top part shows velocity power spectral density for noise, earthquakes (EQs), and LFEs detected using deep learning. For each region, all traces stem from one reference station (Cascadia - MGCB, Guerrero - MAXE, Nankai - IIDH). Noise example have been extracted outside tremor episodes. Spectra have been calculated from the horizontal components (20 s windows for noise, 11 s windows starting 1 s before the S arrival for events). Thin lines show individual spectra, bold lines median spectra. EQs were selected at a similar distance and depth range to the LFEs. Network averaged spectra are shown in Figure S9. The bottom part shows the development of interevent times during the LFE sequences in the reference catalogs, the deep learning catalogs, and for the unmatched events, i.e., all events from the deep learning catalogs that are not in the reference catalogs. Vertical stripes in the events indicate the occurrence of LFE bursts. For Cascadia, we only visualise the 2005 sequence for simplicity. We visualise all events from the reference catalogs without further declustering, leading to very low interevent times.

smaller catalog in Cascadia, a similarly-sized catalog in Nankai, and a far bigger catalog in Guerrero. However, these might be related to changes in the model confidence values rather than their actual quality as we produced all catalogs with fixed picking thresholds. The cross-regional analysis clearly illustrates that the models can be transferred across regions and recover LFEs from families they have not been trained on.

5 Conclusion

Our analysis shows that deep learning and template matching are complementary in the way they detect LFEs with specific advantages and disadvantages for either method. The biggest strength of deep learning is the flexibility. The model can be applied to additional stations, including temporary stations, shows higher diversity in terms of event families, and can be transferred across regions. In addition, our method directly allows to locate single LFEs, even though with substantial uncertainties in terms of depth. In contrast, template matching requires a predefined set of sources that is difficult to obtain and specific to each region and set of stations. While rigid, this leads to a more sensitive model, as evidenced by higher event counts. Furthermore, it allows template matching to identify LFEs with fewer stations than deep learning. For LFEs, commonly no individual location is performed after template matching, due to the difficulties caused by the low SNR ratio. However, it has been shown that relative locations of individual LFEs can be determined with classical methods as well (Shelly et al., 2009). A promising avenue might be the combination of deep learning and template matching, i.e., using deep learning to identify a diverse set of templates and afterwards use template matching to increase the completeness of the identified families.

Lastly, the deep learning method extends our view of LFEs by detecting previously unidentified sources. Building a comprehensive set of templates for template matching is challenging: the low bandwidth and SNR makes it difficult to distinguish between closely spaced sources, leading to a trade-off between missing sources and redundant templates. In contrast, the deep learning method is source-agnostic, i.e., no selection of sources needs to be performed for detecting individual events. Such a source-agnostic view is necessary to perform unbiased subsequent analysis that requires a complete view of LFE sources, such as estimates of slow slip. In addition, the fact that the model can be transferred across regions shows that LFEs have universal, region-independent properties similar to earthquakes. Given our results, we expect that deep learning methods will allow to map LFEs across world regions with high consistency and diversity.

Acknowledgements

We thank David Shelly for sharing his data on the San Andreas LFE catalog and for his insightful comments on the manuscript. We thank Jack Muir and an anonymous reviewer for their comments that helped improve the manuscript. We thank Albanne Lecointre for help with processing the FNet data. This work has been supported by MIAI@Grenoble Alpes (ANR-19-P3IA-0003). This work has been partially funded by the European Union under the grant agreement n°101104996 (“DECODE”) and the ERC CoG 865963 DEEP-trigger. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them. This work was granted access to the HPC/AI resources of IDRIS under the allocation 2022-AD011012345R1 made by GENCI.

Data and code availability

The stack data sets for Cascadia, San Andreas, and Guerrero are available through SeisBench (https://github.com/seisbench/seisbench, https://doi.org/10.5281/zenodo.5568812) from version v0.7 onwards. The model including pretrained weights is available through SeisBench starting from the same version. The INSTANCE noise dataset is available through SeisBench and at https://doi.org/10.13127/instance. The PNSN tremor catalog is available at https://pnsn.org/tremor/. The FNet data is available from the NIED at https://www.fnet.bosai.go.jp/. We use waveforms from the BK (Northern California Earthquake Data Center, 2014a), BP (Northern California Earthquake Data Center, 2014b), CI (California Institute of Technology and United States Geological Survey Pasadena, 1926), CN (Natural Resources Canada, 1975), C8 (Natural Resources Canada, 2002), NC (USGS Menlo Park, 1966), PB (UNAVCO, United States of America, 2004), PO (Geological Survey of Canada, 2000), TA (IRIS Transportable Array, 2003), and TO (MASE deployment; Caltech, 2007) networks. Waveforms were accessed through the IRIS and NCEDC FDSN webservices, and from the PNSN.

Competing interests

The authors have no competing interests.

References

Armbruster, J. G., Kim, W.-Y., and Rubin, A. M. Accurate Tremor Locations from Coherent S and P Waves. *Journal of Geophysical Research: Solid Earth*, 119(6):5000–5013, 2014. doi: 10.1002/2014JB011133.

Bombardier, M., Dosso, S. E., Cassidy, J. F., and Kao, H. Tackling the Challenges of Tectonic Tremor Localization Using Differential Traveltimes and Bayesian Inversion. *Geophysical Journal International*, 234(1):479–493, July 2023. doi: 10.1093/gji/ggad086.

Bostock, M. G., Thomas, A. M., Savard, G., Chuang, L., and Rubin, A. M. Magnitudes and Moment-Duration Scaling of Low-Frequency Earthquakes beneath Southern Vancouver Island. *Journal of Geophysical Research: Solid Earth*, 120(9):6329–6350, 2015. doi: 10.1002/2015JB012195.

California Institute of Technology and United States Geological Survey Pasadena. Southern California Seismic Network, 1926. doi: 10.7914/SN/CI.

Caltech. Meso America Subduction Experiment, 2007. doi: 10.7909/C3RN35SP.

Costantino, G., Giffard-Roisin, S., Radiguet, M., Dalla Mura, M., Marsan, D., and Socquet, A. Multi-station deep learning on geodetic time series detects slow slip events in Cascadia. *Communications Earth & Environment*, 4(1):435, 2023. doi: 10.1038/s43247-023-01107-7.

Cruz-Atienza, V. M., Tago, J., Villafuerte, C., Wei, M., Garza-Girón, R., Dominguez, L. A., Kostoglodov, V., Nishimura, T., Franco, S., Real, J., et al. Short-term interaction between silent and devastating earthquakes in Mexico. *Nature communications*, 12(1): 2171, 2021. doi: 10.1038/s41467-021-22326-6.

Dragert, H., Wang, K., and James, T. S. A silent slip event on the deeper Cascadia subduction interface. *Science*, 292(5521): 1525–1528, 2001. doi: 10.1126/science.1060152.

Frank, W. B. and Brodsky, E. E. Daily measurement of slow slip from low-frequency earthquakes is consistent with ordinary earthquake scaling. *Science advances*, 5(10):eaaw9386, 2019. doi: 10.1126/sciadv.aaw9386.

Frank, W. B. and Shapiro, N. M. Automatic Detection of Low-Frequency Earthquakes (LFEs) Based on a Beamformed Network Response. *Geophysical Journal International*, 197(2): 1215–1223, May 2014. doi: 10.1093/gji/ggu058.

Frank, W. B., Shapiro, N. M., Husker, A. L., Kostoglodov, V., Romanenko, A., and Campillo, M. Using Systematically Characterized Low-Frequency Earthquakes as a Fault Probe in Guerrero, Mexico. *Journal of Geophysical Research: Solid Earth*, 119(10): 7686–7700, 2014. doi: 10.1002/2014JB011457.

Geological Survey of Canada. Portable Observatories for Lithospheric Analysis and Research Investigating Seismicity (POLARIS), 2000. https://fdsn.org/networks/detail/PO/.

Gombert, B. and Hawthorne, J. C. Rapid Tremor Migration During Few Minute-Long Slow Earthquakes in Cascadia. *Journal of Geophysical Research: Solid Earth*, 128(2):e2022JB025034, 2023. doi: 10.1029/2022JB025034.

González-Vidal, D., Moreno, M., Sippl, C., Baez, J. C., Ortega-Culaciati, F., Lange, D., Tilmann, F., Socquet, A., Bolte, J., Hormazabal, J., et al. Relation between oceanic plate structure, patterns of interplate locking and microseismicity in the 1922 Atacama seismic gap. *Geophysical Research Letters*, 50(15): e2023GL103565, 2023. doi: 10.1029/2023GL103565.

Ide, S., Beroza, G. C., Shelly, D. R., and Uchide, T. A Scaling Law for Slow Earthquakes. *Nature*, 447(7140):76–79, May 2007a. doi: 10.1038/nature05780.

Ide, S., Shelly, D. R., and Beroza, G. C. Mechanism of deep low frequency earthquakes: Further evidence that deep non-volcanic tremor is generated by shear slip on the plate interface. *Geophysical Research Letters*, 34(3), 2007b. doi: 10.1029/2006GL028890.

Imanishi, K., Uchide, T., and Takeda, N. Determination of focal mechanisms of nonvolcanic tremor using S wave polarization data corrected for the effects of anisotropy. *Geophysical Research Letters*, 43(2):611–619, 2016. doi: 10.1002/2015GL067249.

IRIS Transportable Array. USArray Transportable Array, 2003. doi: 10.7914/SN/TA.

Japan Meteorological Agency. The Seismological Bulletin of Japan, 2023. http://www.data.jma.go.jp/svd/eqev/data/bulletin/index_e.html.

Kao, H., Shan, S.-J., Dragert, H., Rogers, G., Cassidy, J. F., and Ramachandran, K. A Wide Depth Distribution of Seismic Tremors along the Northern Cascadia Margin. *Nature*, 436 (7052):841–844, Aug. 2005. doi: 10.1038/nature03903.

Lin, J.-T., Thomas, A., Bachelot, L., Toomey, D., Searcy, J., and Melgar, D. Detection of Hidden Low-Frequency Earthquakes in Southern Vancouver Island with Deep Learning. *EarthArxiv*, Oct. 2023. doi: 10.31223/x5pm4b.

Lomax, A., Virieux, J., Volant, P., and Berge-Thierry, C. Probabilistic earthquake location in 3D and layered models: Introduction of a Metropolis-Gibbs method and comparison with linear locations. *Advances in seismic event location*, pages 101–134, 2000. doi: 10.1007/978-94-015-9536-0_5.

Lowry, A. R., Larson, K. M., Kostoglodov, V., and Bilham, R. Transient fault slip in Guerrero, southern Mexico. *Geophysical Research Letters*, 28(19):3753–3756, 2001. doi: 10.1029/2001GL013238.

Michel, S., Gualandi, A., and Avouac, J.-P. Interseismic coupling and slow slip events on the Cascadia megathrust. *Pure and Applied Geophysics*, 176:3867–3891, 2019. doi: 10.1007/s00024-018-1991-x.

Michelini, A., Cianetti, S., Gaviano, S., Giunchi, C., Jozinović, D., and Lauciani, V. INSTANCE–the Italian seismic dataset for machine learning. *Earth System Science Data*, 13(12):5509–5544, 2021. doi: 10.5194/essd-13-5509-2021.

Mouchon, C., Frank, W. B., Radiguet, M., Poli, P., and Cotte, N. Subdaily Slow Fault Slip Dynamics Captured by Low-Frequency Earthquakes. *AGU Advances*, 4(4):e2022AV000848, 2023. doi: 10.1029/2022AV000848.

Moutote, L., Itoh, Y., Lengliné, O., Duputel, Z., and Socquet, A. Evidence of a transient aseismic slip driving the 2017 Valparaiso earthquake sequence, from foreshocks to aftershocks. *Journal of Geophysical Research: Solid Earth*, page e2023JB026603, 2023. doi: 10.1029/2023JB026603.

Münchmeyer, J. PyOcto: A high-throughput seismic phase associator. *Seismica*, 3(1), 2024. doi: 10.26443/seismica.v3i1.1130.

Münchmeyer, J., Woollam, J., Rietbrock, A., Tilmann, F., Lange, D., Bornstein, T., Diehl, T., Giunchi, C., Haslinger, F., Jozinović, D., et al. Which picker fits my data? A quantitative evaluation of deep learning based seismic pickers. *Journal of Geophysical Research: Solid Earth*, 127(1):e2021JB023499, 2022. doi: 10.1029/2021JB023499.

Natural Resources Canada. Canadian National Seismograph Network, 1975. doi: 10.7914/SN/CN.

Natural Resources Canada. Canadian Seismic Research Network, 2002. https://fdsn.org/networks/detail/C8/.

Northern California Earthquake Data Center. Berkeley Digital Seismic Network (BDSN), 2014a. doi: 10.7932/BDSN.

Northern California Earthquake Data Center. High Resolution Seismic Network (HRSN), 2014b. doi: 10.7932/HRSN.

Okada, Y., Nishimura, T., Tabei, T., Matsushima, T., and Hirose, H. Development of a detection method for short-term slow slip events using GNSS data and its application to the Nankai subduction zone. *Earth, Planets and Space*, 74(1):1–18, 2022. doi: 10.1186/s40623-022-01576-8.

Ozawa, S., Murakami, M., Kaidzu, M., Tada, T., Sagiya, T., Hatanaka, Y., Yarai, H., and Nishimura, T. Detection and monitoring of ongoing aseismic slip in the Tokai region, central Japan. *Science*, 298(5595):1009–1012, 2002. doi: 10.1126/science.1076780.

Radiguet, M., Perfettini, H., Cotte, N., Gualandi, A., Valette, B., Kostoglodov, V., Lhomme, T., Walpersdorf, A., Cabral Cano, E., and Campillo, M. Triggering of the 2014 M w 7.3 Papanoa earthquake by a slow slip event in Guerrero, Mexico. *Nature Geoscience*, 9 (11):829–833, 2016. doi: 10.1038/ngeo2817.

Ronneberger, O., Fischer, P., and Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, pages 234–241. Springer, 2015. doi: 10.1007/978-3-319-24574-4_28.

Ross, Z. E., Meier, M.-A., Hauksson, E., and Heaton, T. H. Generalized seismic phase detection with deep learning. *Bulletin of the Seismological Society of America*, 108(5A):2894–2901, 2018. doi: 10.1785/0120180080.

Rouet-Leduc, B., Hulbert, C., McBrearty, I. W., and Johnson, P. A. Probing slow earthquakes with deep learning. *Geophysical research letters*, 47(4):e2019GL085870, 2020. doi: 10.1029/2019GL085870.

Royer, A. and Bostock, M. A comparative study of low frequency earthquake templates in northern Cascadia. *Earth and Planetary Science Letters*, 402:247–256, 2014. doi: 10.1016/j.epsl.2013.08.040.

Rubin, A. M. and Armbruster, J. G. Imaging Slow Slip Fronts in Cascadia with High Precision Cross-Station Tremor Locations. *Geochemistry, Geophysics, Geosystems*, 14(12):5371–5392, 2013. doi: 10.1002/2013GC005031.

Savard, G. and Bostock, M. G. Detection and Location of Low-Frequency Earthquakes Using Cross-Station Correlation. *Bulletin of the Seismological Society of America*, 105(4):2128–2142, June 2015. doi: 10.1785/0120140301.

Shelly, D. R. Migrating tremors illuminate complex deformation beneath the seismogenic San Andreas fault. *Nature*, 463(7281): 648–652, 2010. doi: 10.1038/nature08755.

Shelly, D. R. A 15 Year Catalog of More than 1 Million Low-Frequency Earthquakes: Tracking Tremor and Slip along the Deep San Andreas Fault. *Journal of Geophysical Research: Solid Earth*, 122(5):3739–3753, 2017. doi: 10.1002/2017JB014047.

Shelly, D. R., Beroza, G. C., and Ide, S. Non-Volcanic Tremor and Low-Frequency Earthquake Swarms. *Nature*, 446(7133): 305–307, Mar. 2007. doi: 10.1038/nature05666.

Shelly, D. R., Ellsworth, W. L., Ryberg, T., Haberland, C., Fuis, G. S., Murphy, J., Nadeau, R. M., and Bürgmann, R. Precise Location of San Andreas Fault Tremors near Cholame, California Using Seismometer Clusters: Slip on the Deep Extension of the Fault? *Geophysical Research Letters*, 36(1), 2009. doi: 10.1029/2008GL036367.

Socquet, A., Valdes, J. P., Jara, J., Cotton, F., Walpersdorf, A., Cotte, N., Specht, S., Ortega-Culaciati, F., Carrizo, D., and Norabuena, E. An 8 Month Slow Slip Event Triggers Progressive Nucleation of the 2014 Chile Megathrust. *Geophysical Research Letters*, 44 (9):4046–4053, 2017. doi: 10.1002/2017GL073023.

Supino, M., Poiata, N., Festa, G., Vilotte, J. P., Satriano, C., and Obara, K. Self-Similarity of Low-Frequency Earthquakes. *Scientific Reports*, 10(1):6523, Apr. 2020. doi: 10.1038/s41598-020-63584-6.

Tan, Y. J., Waldhauser, F., Ellsworth, W. L., Zhang, M., Zhu, W., Michele, M., Chiaraluce, L., Beroza, G. C., and Segou, M. Machine-learning-based high-resolution earthquake catalog reveals how complex fault structures were activated during the 2016–2017 Central Italy sequence. *The Seismic Record*, 1(1): 11–19, 2021. doi: 10.1785/0320210001.

Thomas, A. M., Inbal, A., Searcy, J., Shelly, D. R., and Bürgmann, R. Identification of low-frequency earthquakes on the San Andreas Fault with deep learning. *Geophysical Research Letters*, 48(13): e2021GL093157, 2021. doi: 10.1029/2021GL093157.

UNAVCO, United States of America. Plate Boundary Observatory Borehole Seismic Network (PBO), 2004. https://fdsn.org/networks/detail/PB/.

USGS Menlo Park. USGS Northern California Seismic Network, 1966. doi: 10.7914/SN/NC.

Wang, Q.-Y., Frank, W. B., Abercrombie, R. E., Obara, K., and Kato, A. What makes low-frequency earthquakes low frequency. *Science Advances*, 9(32):eadh3688, 2023. doi: 10.1126/sciadv.adh3688.

Wech, A. G. Cataloging Tectonic Tremor Energy Radiation in the Cascadia Subduction Zone. *Journal of Geophysical Research: Solid Earth*, 126(10):e2021JB022523, 2021. doi: 10.1029/2021JB022523.

Woollam, J., Münchmeyer, J., Tilmann, F., Rietbrock, A., Lange, D., Bornstein, T., Diehl, T., Giunchi, C., Haslinger, F., Jozinović, D., et al. SeisBench—A toolbox for machine learning in seismology. *Seismological Research Letters*, 93(3):1695–1709, 2022. doi: 10.1785/0220210324.

Zhu, W. and Beroza, G. C. PhaseNet: A deep-neural-network-based seismic arrival-time picking method. *Geophysical Journal International*, 216(1):261–273, 2019. doi: 10.1093/gji/ggy423.

doi:10.26443/seismica.v3i1.1166

SEISMICA

Monitoring urban construction and quarry blasts with low-cost seismic sensors and deep learning tools in the city of Oslo, Norway

Andreas Köhler * [1,2], **Erik Myklebust** [1], **Anna Maria Dichiarante** [1], **Volker Oye** [1]

[1]NORSAR, Kjeller, Norway, [2]UiT - The Arctic University of Norway, Tromsø, Norway

Author contributions: *Conceptualization*: Andreas Köhler. *Methodology*: Erik Myklebust, Andreas Köhler. *Software*: Erik Myklebust, Andreas Köhler. *Formal Analysis*: Andreas Köhler. *Writing - Original draft*: Andreas Köhler. *Writing - Review & Editing*: Anna Maria Dichiarante. *Visualization*: Anna Maria Dichiarante. *Project administration*: Volker Oye. *Funding acquisition*: Volker Oye.

Abstract The aim of this study is to collect information about events in the city of Oslo, Norway, that produce a seismic signature. In particular, we focus on blasts from the ongoing construction of tunnels and under-ground water storage facilities under populated areas in Oslo. We use seismic data recorded simultaneously on up to 11 Raspberry Shake sensors deployed between 2021 and 2023 to quickly detect, locate, and classify urban seismic events. We present a deep learning approach to first identify rare events and then to build an automatic classifier from those templates. For the first step, we employ an outlier detection method using auto-encoders trained on continuous background noise. We detect events using an STA/LTA trigger and apply the auto-encoder to those. Badly reconstructed signals are identified as outliers and subsequently located using their surface wave (Rg) signatures on the seismic network. In a second step, we train a supervised classifier using a Convolutional Neural Network to detect events similar to the identified blast signals. Our results show that up to 87% of about 1,900 confirmed blasts are detected and locatable in challenging background noise conditions. We demonstrate that a city can be monitored automatically and continuously for explosion events, which allows implementing an alert system for future smart city solutions.

Production Editor:
Gareth Funning
Handling Editor:
Lise Retailleau
Copy & Layout Editor:
Hannah F. Mark

Signed reviewer(s):
René Steinmann

Received:
December 15, 2023
Accepted:
May 27, 2024
Published:
June 17, 2024

Non-technical summary Monitoring infrastructures and operations in cities relies on different kinds of sensors providing information for local authorities and the general public. In this study we collect information about events in the city of Oslo, Norway, that produce ground shaking. We focus on blasts from the ongoing construction of tunnels and under-ground storage facilities under populated areas in Oslo. We use data from senors in the city, deployed between 2021 and 2023 for example in schools, to identify these blasts by means of machine learning methods. We are able to detect up to 87% of about 1,900 confirmed blasts.

1 Introduction

Global estimates for future growth indicate that the population of cities will continue to increase (Brockerhoff, 1999). This growth has caused many cities to upgrade their infrastructures and to embrace the vision of a "smart-city" (McKinsey, 2018). Data collection through different types of sensors represents the base layer for such solutions. Large data sets are being produced and need to be automatically processed so that relevant information can be extracted and transferred to local authorities and the general public to facilitate decisions and to optimize the performance of cities in areas such as transport, safety and supply of water and energy (Fischer et al., 2013; Chang et al., 2014; Al Nuaimi et al., 2015).

Integrating seismic data into the data collection of such systems is currently not a common and widespread approach, although the potential of urban seismology using seismometers or Distributed Acoustic Sensing (DAS) has already been recognized in previous studies (Ritter et al., 2005; Díaz et al., 2017; Spica et al., 2020). To date, this approach is routinely used mainly for earthquake early warning and fast response in urban areas with substantial seismic hazard (Kong et al., 2016), or for monitoring geothermal or other reservoirs in proximity to cities (Kraft et al., 2009; Hillers et al., 2020; Fiori et al., 2023). Advantages of using seismic data to monitor other urban activities compared to, for example, optical and acoustic systems include better compliance with General Data Protection Regulations (GDPR) (Zhang et al., 2017), efficient propagation of signals in the ground, independence of visibility, and in general a new type of sensor data not provided by the other methods.

This study focuses on the city of Oslo, Norway, addressing common needs of two departments of the municipality, i.e., the Agency for Emergency Planning and the Water and Sewage Department. The Agency for Emergency Planning is interested in obtaining quick information about any kind of unusual event that pro-

*Corresponding author: andreas.kohler@norsar.no

duces a seismic signature, e.g., explosions or sudden mass movements, to facilitate fast emergency response. An example of such an event was the bombing of a government building in the city center of Oslo during the terrorist attack on 22 July 2011 which was recorded on seismometers in and around Oslo (Bergen University, 2012). The Water and Sewage Department is concerned with monitoring ongoing construction activity to secure the freshwater supply of the city of Oslo. The construction of tunnels and under-ground water storage facilities under populated areas started in 2021 and is planned to be finished in 2028. Furthermore, due to population growth in Oslo, public transport infrastructure is currently extended, i.e., new metro tunnels are being constructed below or close to residential areas. Finally, a tunnel for a main electric power line under the city has been under construction since 2023. All these construction activities are accommodated by blasts which are partly felt by citizens, which have raised concerns in the population during a few documented incidents when the explosion yield was higher than anticipated.

Explosion monitoring with seismic sensors is a well-established technique for observing mining and quarry activities on a regional scale (Gibbons and Ringdal, 2006) or for verifying the Comprehensive Nuclear Test Ban Treaty (CTBTO) on a global scale (Kalinowski and Mialle, 2021). More recently it has also been used for identifying military attacks (Dando et al., 2023). A challenge with pursuing such an approach in urban areas on a very local scale and preferably in real-time, is the presence of a multitude of other seismic sources and high background noise levels. Such complex records require advanced processing methods which may be found in machine (ML) or deep learning (DL), fields which have made great advances within seismology in recent years (Kong et al., 2019; Bergen et al., 2019; Mousavi et al., 2019; Mousavi and Beroza, 2023; Zhu and Beroza, 2018; Mousavi et al., 2020; Provost et al., 2017).

In this context, there are two main possible approaches we can pursue: (1) Identification of so far unidentified seismic events of interest in an unsupervised manner or (2) using a sufficiently large number of already identified events of interest to train a classifier in a supervised manner. Approach (1) will be required in most cases as an initial step for urban monitoring purposes. It can be further divided into clustering, where the outcome are groups of signals or time windows of potential interest, or outlier detection, where the target is only a particular group of infrequent events. Clustering can be either done by automatically grouping the continuous seismic records (Köhler et al., 2010; Johnson et al., 2020; Chamarczuk et al., 2020; Seydoux et al., 2020; Steinmann et al., 2022a,b) or by grouping pre-detected transient signals (Sick et al., 2015; Jenkins et al., 2021). The features a clustering algorithm utilizes must be either extracted beforehand (e.g., Köhler et al., 2010) or are extracted automatically by a DL architecture (e.g., Mousavi et al., 2019). In a broader sense, simple non-machine learning methods, such as the well-known Short-Term Average over Long-Term Average (STA/LTA) trigger or trigger algorithms based on other characteristic functions of the seismic waveforms (kurtosis, spectral amplitudes in different bands, etc.), may be considered to belong to approach (1). They can be used directly or combined with clustering for outlier detection. Hence, we can consider the STA/LTA method to be the baseline which ML or DL methods must outperform. In other words, under certain conditions STA/LTA may still be the most efficient way to detect events of interest.

In this study, we use passive seismic records acquired with the objective to quickly detect, classify and locate urban seismic events, particularly blasts. The use cases for detecting those events in near real-time include, but are not limited to, quickly informing the public in case of ground shaking felt by citizens or quickly identifying large blasts from construction works or attacks that can impact public safety and infrastructure integrity due to potential damage caused to structures (Shallan et al., 2014; Dowding, 2016; Naveen et al., 2021) or mobilization of unstable ground (Bouchard et al., 2018). For this purpose, a seismic network of low-cost sensors was deployed in target areas within the city of Oslo from spring 2021 onwards. We present a DL approach to first identify target events and then, if target events are sufficient in number, to build an automatic classifier from those templates. For the first step, we suggest an outlier detection method for automatic identification of rare events. These events are then located using their short-period fundamental-mode Rayleigh wave (Rg) signatures on the seismic network by means of stacking the observed travel-time corrected waveform envelopes. We then identify blasts inside and close to the city limits of Oslo and use them to train a supervised deep learning classification method to detect more of these events missed by the outlier detector.

2 The seismic network

We deployed three-component Raspberry Shake 3D sensors (Nugent, 2018) at different locations within the city of Oslo (Fig. 1, Table 1). The network was extended gradually starting in May 2021, with up to 11 stations recording simultaneously from June 2022 to July 2023. The sensors were connected to mobile network modems for real-time data transmission and remote maintenance. GPS antennas were deployed where possible. However, we found that the timing provided through NTP (Network Time Protocol) was sufficient at a few sites where free view to the sky could not be established. Sensor locations were in the basement of private businesses, private houses, and public school buildings. The first batch of sensors (ALNN1-4, ALNN7) was deployed with a dense layout in an area with quick clay in the sub-surface in the Eastern part of Oslo (Alna area) to allow for near-surface structural monitoring using ambient noise and detection of possible ground movements, a task which will be described in a future study. ALNN2 was removed after a few months in November 2021 due to construction activities in the host building. More sensors were deployed to the North of that area (ALNN5, ALNN6, ALNN8) and towards the city center and the Western part of Oslo (EKBG1, OSLN1-OSLN5).

The latter batch of sensors was located closer to the area of activity related to the construction of tunnels and an underground cavity for freshwater storage (Fig. 1), where the excavation is mostly done by blasting. All data are recorded continuously with a 100 Hz sampling rate, and the corner frequency of the sensors is at about 0.5 Hz. In addition, we use two permanent seismic stations equipped with broadband seismometers, one located on the university campus (OSL, part of Norwegian National Seismic Network) (Ottemöller et al., 2018, 2021) and the other one outside the city, to the Southeast of Oslo (OFNS2, not on map in Fig. 1).

3 Methods

Our aim is to detect rare or unusual seismic events observed on the deployed network using an outlier detection method. In contrast to a standard STA/LTA trigger approach, we do not want to simply detect all transient signals in the data stream. Frequently and regularly occurring urban events or noise bursts only recorded at single stations are not the focus of this study, although for example traffic monitoring with seismic data might be another topic of interest in urban seismology. Outlier events in our definition are singular or repeating events, but the latter not dominating the record, i.e., occurring not more than a few times per day. Hence, here we do not pursue a full clustering of all occurring signal and noise classes using ML, DL or other big data methods as done in previous studies (Köhler et al., 2010; Yoon et al., 2015; Seydoux et al., 2020; Steinmann et al., 2022a). However, it should be noted that clustering can be used for outlier detection. It would require to identify the event cluster of interest, i.e., the rare events. However, we decided to not pursue this approach further since we want to avoid the manual step to identify the outlier cluster. Furthermore, rare events may not necessarily be caught up in a distinct cluster.

Our workflow begins with identifying and collecting these repeating outlier events based on their origin locations and, if a sufficient number of observations has been collected, building a supervised classifier with labeled training data to more reliably detect those events in continuous data. Doing so, we have an outlier detection method available shortly after the start of the measurements which is flexible enough to pick up new events, while the supervised classifier can be gradually enhanced during the course of the seismic deployment by training it with newly identified events.

3.1 Outlier detection

Auto-encoder neural networks are popular methods for dimensionality reduction (Wang et al., 2016) and to identify anomalies or outliers in time series data (Yin et al., 2020; Thill et al., 2021). The idea is to use a Convolutional Neural Network (CNN) to reduce the dimension of the input time series, here three-component seismic waveforms with T time samples per component, using a series of convolutional layers or filters, and then use the resulting latent features to reconstruct the signal with a mirrored model neural network structure (Fig. 2). In seismology this approach has been mostly adapted for data compression and interpolation (Navarro et al., 2019; Zheng and Zhang, 2020; Nuha et al., 2020). Furthermore, Valentine and Trampert (2012) highlighted the potential of auto-encoders for various applications in seismology. Mousavi et al. (2019) used an auto-encoder model to extract features suitable for unsupervised clustering. If data compression is the goal, the number of latent features should be low. However, our objective here is not primarily to compress data, and therefore we tested different dimensions from no compression at all, i.e., number of latent features is equal to $3T$, down to a latent dimension of T. Our final choice with best performance is a number of $2T$ latent features, i.e., a data compression by 33% (see supplementary Figure S1 for a comparison).

If the auto-encoder is trained using a continuous (unlabeled) seismic record, which is representative for a particular station, waveforms of regularly occurring signals and background noise should be reconstructed well by the model. If a signal is not reconstructed well enough, it can be considered to be an outlier. This approach has some relation to an auto-regressive model, which predicts a time series based on previously observed data and which is well known in seismology for its ability to detect the onset of seismic arrivals (Leonard and Kennett, 1999). However, similar to the STA/LTA method, an auto-regressive model is sensitive to all (including frequently occurring) transient signals with different characteristics than the background noise, a property which is not desired in our case.

We train auto-encoder models for single stations using the vertical and both horizontal components. Here, we use two stations with comparable low background noise levels as trigger stations: EKBG1 southeast of the city center and OSLN2 to the west of the city center (Fig. 1). The auto-encoder is trained for the two sites and then applied to time windows including STA/LTA detections obtained from the continuous data. Doing so, we aim to select only those STA/LTA triggered signals that can potentially be of interest. A future extension would ideally include outlier triggers on all stations. However, since in this study we are only interested in locatable events observed simultaneously on several stations of our network, we found it to be sufficient to trigger only on these two stations, since all locatable events are observed on at least one of these.

The auto-encoder input differs slightly for both stations, and also the selection of training data is done in a different manner. For OSLN2 the training data is a continuous record of eight consecutive days (02/06/2022–09/06/2022) band-pass filtered between 0.3–12.5 Hz. The size of the input time segment fed into the auto-encoder is $T = 512$ samples (see Fig. 2) for each component. For EKBG1 we use a higher number of samples ($T = 1024$), partly because this record visually appeared a bit more complex (frequent transients). A band-pass filter between 0.3 and 25 Hz is used to potentially also capture outliers with higher frequency content. The training data for EKBG1 are 90 time periods of continuous data of 6 hours' duration each, selected between 02/11/2021 and 14/03/2022. The

Name	Longitude	Latitude	Recording time	Location
ALNN1	10.8582	59.9336	08.06.2021-30.09.2023	Alfaset graveyard
ALNN2	10.8497	59.9282	09.06.2021-15.11.2021	private business
ALNN3	10.8452	59.9300	21.05.2021-30.09.2023	private business
ALNN4	10.8480	59.9314	09.06.2021-30.09.2022	private business
ALNN5	10.8336	59.9409	25.09.2021-30.09.2023	Linderud public school
ALNN6	10.8353	59.9405	29.09.2021-30.09.2023	Linderud public school
ALNN7	10.8464	59.9302	30.09.2021-30.09.2023	private business
ALNN8	10.8373	59.9411	15.11.2021-22.01.2023	Bjerke public school
EKBG1	10.7581	59.8974	03.11.2021-30.09.2023	Kongshavn public school
OSLN1	10.7694	59.9552	27.04.2022-30.09.2023	private house
OSLN2	10.7062	59.9415	01.06.2022-30.09.2023	Vinderen public school
OSLN3	10.7328	59.9425	01.06.2022-13.09.2023	Ullevål public school
OSLN4	10.6548	59.9415	24.10.2022-24.06.2023	Hovseter public school
OSLN5	10.7670	59.9650	10.06.2023-30.09.2023	private house
OSL	10.7227	59.9372	permanent NNSN station	
OFSN2	10.9108	59.8401	permanent NORSAR station	

Table 1 Seismic stations used in the study.

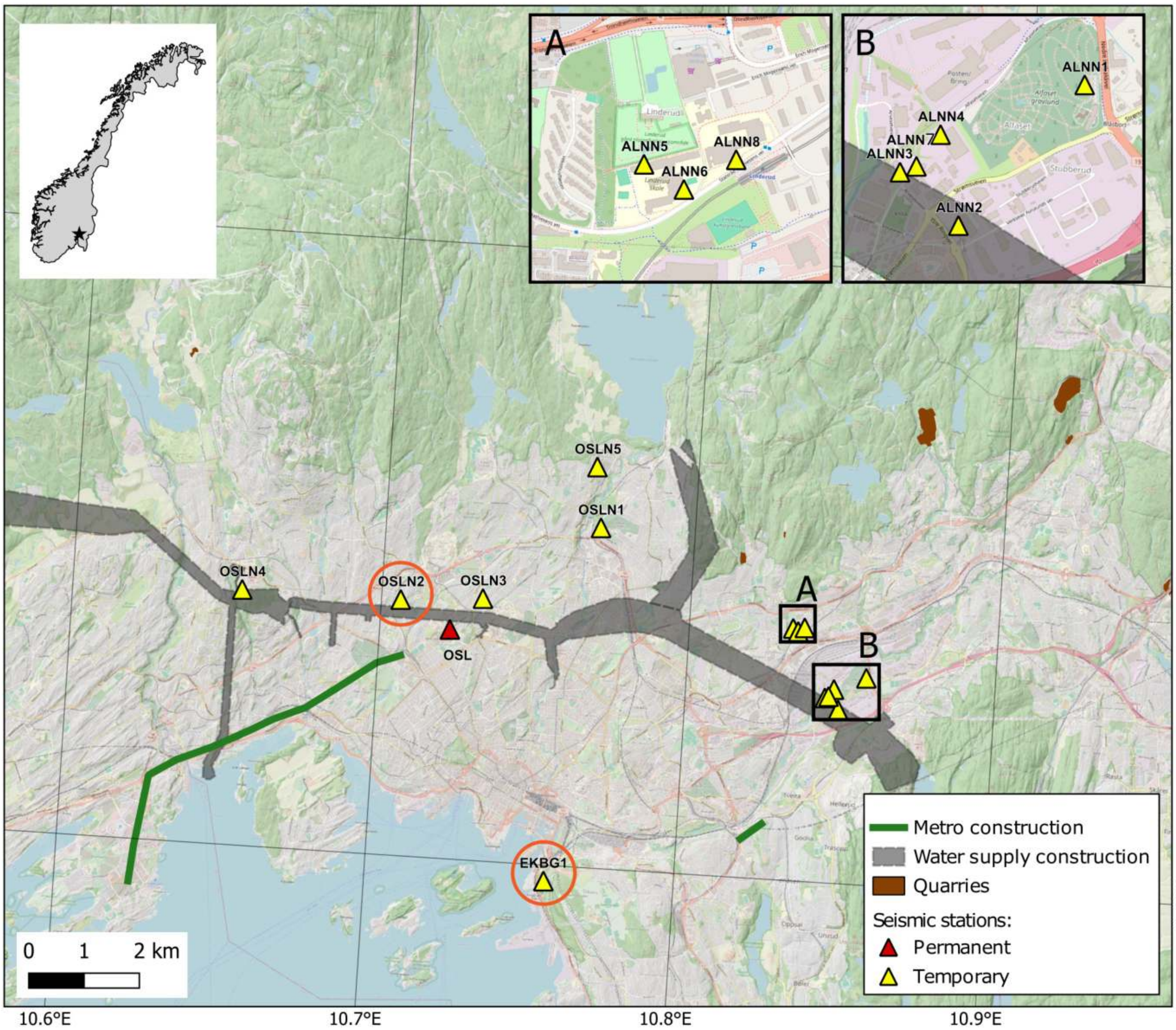


Figure 1 Map of the city of Oslo (OpenStreetMap contributors, 2017) draped on the Digital Elevation Model (DEM) at 10 m resolution, with infrastructure, potential seismic source areas and seismic stations. Map location is shown on the top left inset. (A, B) Close-ups of two areas in Eastern Oslo with denser seismic deployment. Stations which were used for STA/LTA triggering and outlier detection are marked with orange circles.

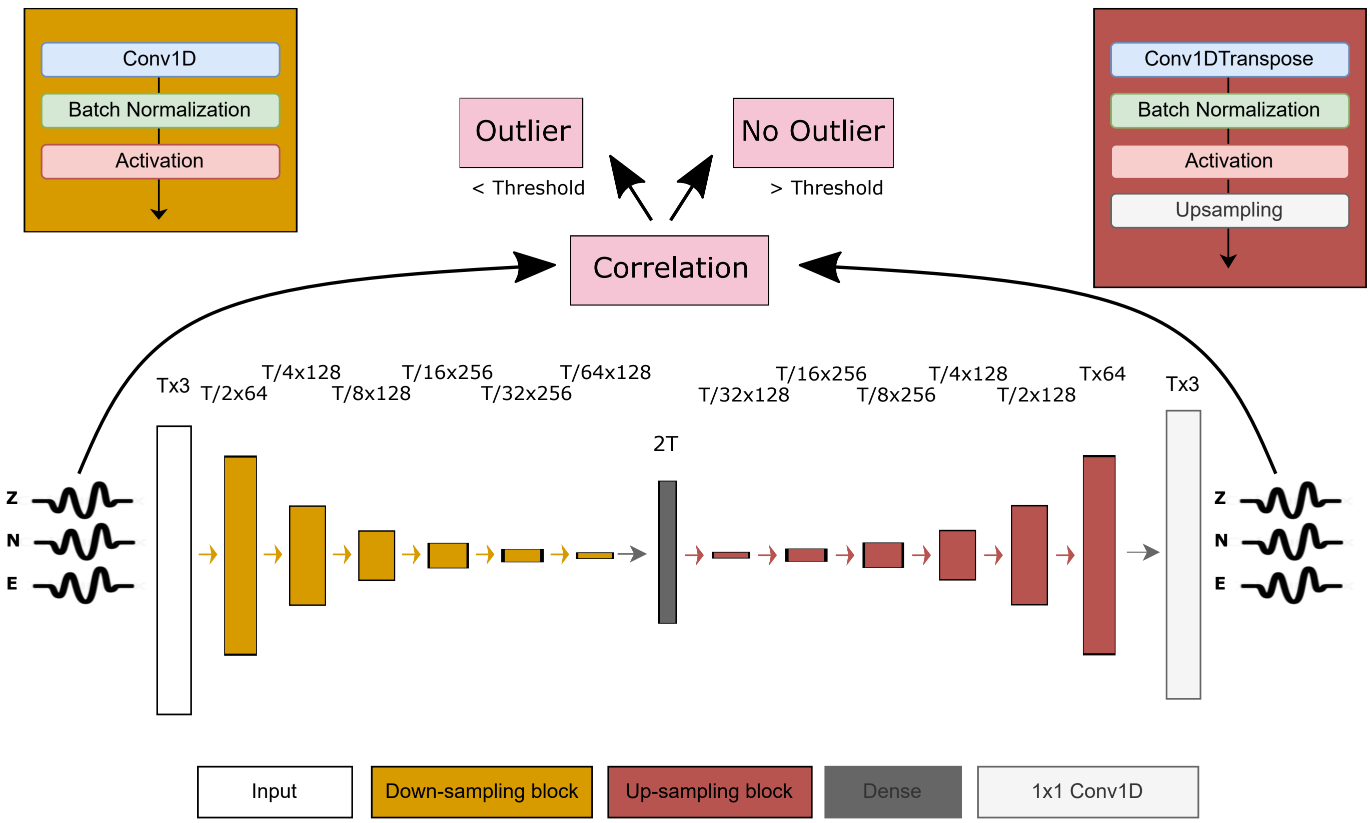


Figure 2 Auto-encoder architecture consisting of several down-sampling and up-sampling blocks shown in detail above the neural network. T represent the number of time-steps of one component. After down-sampling, the output is flattened and a dense layer is used to create the latent dimension ($2T$). The latent dimension is reshaped and used as input to the series of up-sampling blocks. Finally, a single convolutional layer is used to construct the output of the network. Each convolutional layer uses between 64 and 256 filters with a kernel size of 7. Input and output signals are cross-correlated to detect badly reconstructed events with low correlation coefficients, i.e., outliers.

motivation for this selection was to exclude visually detected blast signals from the training. This was achieved by manually screening a selection of days, from which the 6 hour-long time windows were then chosen. We did not pursue the same approach for OSLN2, which was implemented later, since we found that keeping the rare outlier events in the training data did not impair detection performance. In general, we found it to be more important that representative noise records are included (i.e., day and night, weekday and weekend), rather than making sure that outlier events of interest are excluded from the auto-encoder training. All waveform time windows are normalized with minimum and maximum amplitude before being fed into the auto-encoder.

We apply an STA/LTA detector with a low threshold (STA length = 0.5 s, LTA length = 10 s, STA/LTA threshold = 4) to continuous three-component data, and if all three components exhibit a coincident trigger, a detection is declared. We then apply the auto-encoder method to a time window around each detection (same duration as training data). One way of evaluating the quality of the auto-encoder signal reconstruction, i.e., to identify an outlier signal, is to compute the normalized correlation coefficient of the original and reconstructed seismic traces (Fig. 2). We also tested using the reconstruction loss, i.e., RMS value of the observed and reconstructed waveforms, for outlier detection. However, we found the correlation value to be more suitable to identify outlier events which are mostly blasts in our case. A comparison of RMS and correlation values of all STA/LTA detections and confirmed blast signals at station OSLN2 is provided in the supplementary Figure S4, showing that the RMS is not a good discriminant for blasts in our case.

Figure 3 shows examples of one outlier event and three STA/LTA detections not recognized as outliers at each station (OSLN2 (a-d) and EKBG1 (e-h)). The correlation between original and reconstructed traces for most data is in general very high, i.e., larger than about 0.8. Using a lower/higher latent dimension would compress the seismic data more/less, which would increase/decrease the construction loss and decrease/increase the correlation. We found the current dimension of the auto-encoder to be optimal for our task. However, it must be tuned for each new data set. Figure 3a and e show seismic signals which were later confirmed as blasts. Particularly, the vertical component waveforms are not well reconstructed. Hence, both exhibit comparably low correlation coefficients (0.6 and 0.84) in relation to the other STA/LTA detected transient signals. The latter are manually identified as regularly occurring noise bursts, including signals originating from very close to the sensors.

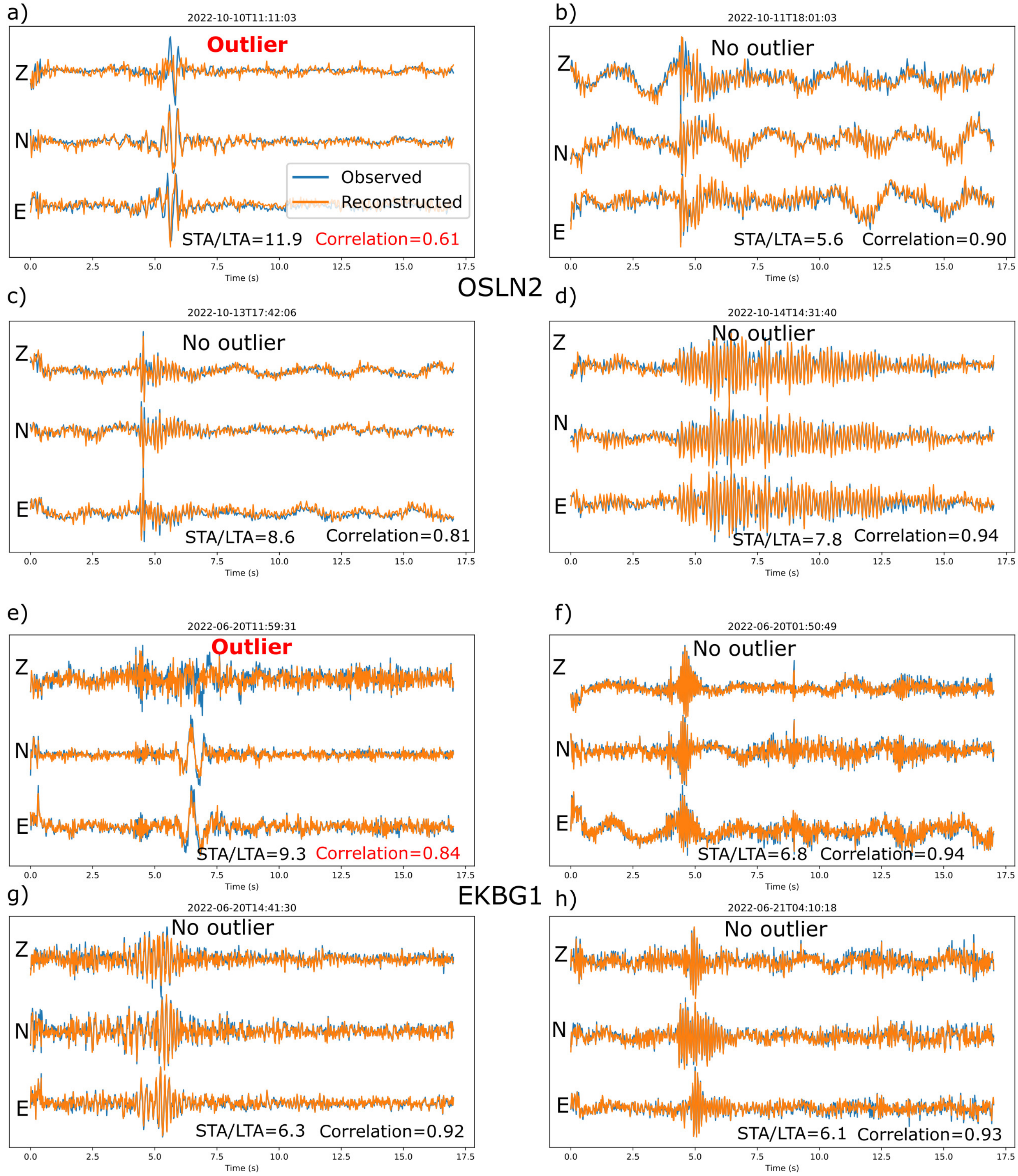


Figure 3 Comparison of recorded three-component waveforms (blue) and waveforms predicted by auto-encoder (orange) for station OSLN2 and EKBG1. STA/LTA value and correlation coefficient between traces are given in each panel. Two detected outlier events, marked red in a) and e), are confirmed blast signals. The other transient signals triggered by the STA/LTA method were not detected as outliers due to higher correlation.

3.2 Locations based on Rg waves

Once an outlier is detected at station OSLN2 or EKBG1, the remaining stations are used in addition to attempt an automatic location. P and S wave arrivals would be needed for traditional event location based on onset time readings, but are not observed for the majority of events due to high noise levels. However, we found the Rg wave, which is a short-period Rayleigh wave in the Earth crust typically observed for seismic sources close to the surface, to be well recorded over the entire network between 0.8 and 2 Hz. In order to use Rg for event location, we first compute envelopes of the band-pass filtered vertical component data and discard stations with Signal-to-Noise Ratios (SNR) below 7.0 (OSLN2) and 6.2 (EKBG1). If a minimum of four stations are left, we perform a 2D gridsearch to find the event source location maximizing the stacked time-shifted envelopes of all stations, an approach similar to stacking and migration methods developed for seismic event localization (Gharti et al., 2010). For Rg travel time computation we assume a velocity of 2.0 km/s which we found to fit observed Rg waves with plane wave fronts travelling over the network from a known source at a large distance. The extent of the gridsearch is from 10.55 to 11.0 degrees longitude and 59.86 to 60.0 degrees latitude. The step width is 0.01 degrees in longitude and 0.005 degrees in latitude. In addition to maximizing envelope stacks, we estimate the Rg back-azimuth from three-component data for the stations above the SNR threshold. This is done by finding the rotation angle maximizing the amplitude on the radial component. The 180 degree ambiguity is avoided by selecting the direction whose correlation coefficient between the vertical and Hilbert-transformed radial component of the Rg wave is positive. The weighted back-azimuth residual of each location grid point is subtracted from the stacking amplitude. The grid point maximizing this value is taken as the source location. Based on the location, we assign to each confirmed and locatable outlier a label corresponding to different construction areas in the city of Oslo.

3.3 CNNs for blast classification

The supervised classifier is a Convolutional Neural Network (CNN) which takes three-component waveform data of a single seismic station as input (Fig. 4). The method uses the well-established AlexNet architecture (Krizhevsky et al., 2012) and is loosely based on the model we used in Köhler et al. (2022) to classify calving events in the Arctic. We train a two-class model distinguishing STA/LTA detections of blasts in Oslo and all other detections (noise and other events). Here, we only use station OSLN2 to train and test the classifier. The model consists of a layer to randomly crop the input waveforms, five convolutional layers with batch normalization and max-pooling, and finally two dense layers which process the flattened output of the convolutional layers and generate the output probabilities. We use 26 s as input time window duration around each blast which is cut to 22 s by random cropping. For the noise class we use a time window of waveform data of the same duration before each blast detection such that the classes are balanced. The hyper-parameters controlling size of convolutional filters and type of pooling are tuned with KerasTuner (Chollet et al., 2015). In the tuning process, the number of filters in each of the five convolutional layers was kept constant. KerasTuner uses ranges of hyper-parameters (filter length between 3 and 49) and different options (max vs. average pooling) as input and searches the parameter space to optimize the classification accuracy. The final hyper-parameters are shown in Figure 4.

For the final classifier, we use a stratified 5-fold cross validation, i.e., five different models are trained, each using 80% of the shuffled data (confirmed blasts from the outlier detection and the noise class examples) for training and 20% for validation. When applying the classifier for prediction, the averaged probabilities for blast and not blast of these five models are used. Finally, we have to set a probability threshold for detecting a blast. We can either use a threshold of 0.5, i.e., select the winning class, or require a higher confidence for a blast to be detected using a higher threshold.

3.4 Reference blast detections based on STA/LTA method

For evaluating the outlier detection method and the blast classifier, we would need complete ground-truth data about blast occurrence in the city of Oslo, which turned out to be difficult to obtain. Alternatively, we can visually screen all potential seismic events observed on the network. This will not allow us to assess the network's detection sensitivity, but rather the detection method's ability to recognize all recorded blasts. Since our methods use STA/LTA detections as input, we create our reference event catalog by processing all these STA/LTA detections in the same way we process the outliers, i.e., attempt a Rg wave based localization and label the recognized blast signals in the different parts of the city. This resulted in 1,870 blasts located within the study area between November 2021 and October 2023.

We use the recall and precision metrics to evaluate all deep learning models with respect to this data set. We want to achieve a high recall (high number of recognized blasts) and high precision (low number of false triggers). In contrast to a conventional event detector, the decision on what is a false and what is a true positive is a moving target for an outlier detector. Additional events not being part of the reference data could still be events of interest. Nevertheless, we can still use the recall-precision metrics as a proxy to compare model performance relative to each other. To decide on decision thresholds for the outlier detector (correlation coefficient) and blast classifier (blast probability), we evaluate recall-precision curves provided in the supplementary material. We compute the performance metrics before event localization since we want to include unlocatable events.

4 Results of outlier event detection

After visual inspection of many signals detected on multiple days as well as evaluating recall and precision, we finally set the outlier detection threshold to 0.86 for EKBG1 and 0.78 for OSLN2. Supplementary Figure S1 shows recall-precision curves for outlier detectors at OSLN2. The optimal detection threshold corresponds to correlation coefficients producing a recall-precision point closest to the upper-left corner. Note that precision will increase after applying the automatic location procedure because false alarms are usually unlocatable events. The distribution of correlation coefficients at stations OSLN2 and EKBG1 in relation to the

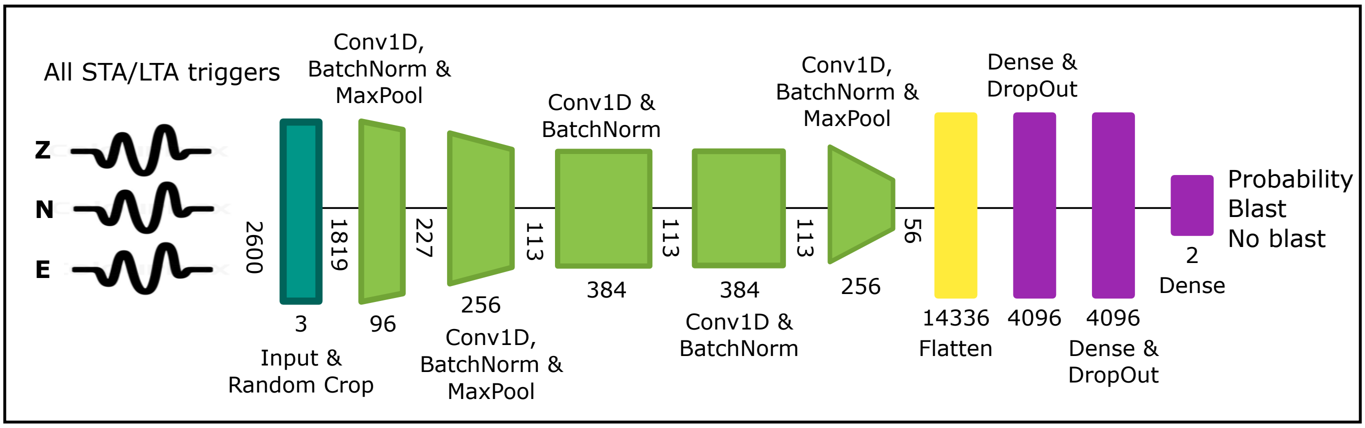


Figure 4 Architecture of the deep Neural Network of the blast classifier.

chosen thresholds are provided in supplementary Figure S5.

Figures 5, 6 and 7 show maps of stacked envelope amplitudes and the best location estimates for several blast signals detected as outliers. The events in Figure 5a–c are blasts related to the construction of underground water tunnels and storage facilities at the Stubberud, Oset, and Huseby sites (see also Fig. 1). Figure 6 shows a blast at the Losby quarry East of Oslo, a blast at the metro tunnel construction site at Bryn, and a blast from a construction site for a new electric power line tunnel in Sogn. For the latter we have a ground-truth confirmation of location and blasting time. Figure 7 shows three blast signals for which we also have ground-truth times and locations. They originate from the construction of the new metro tunnel for the Fornebu line in the West of Oslo, about 3 km south of the underground water storage construction site. The precise locations were provided to us and the public with an uncertainty of about 100 m.

The spatial distribution of stacking amplitudes shows as expected that the resolution strongly depends on the event location. Outside of the network, resolution is poorer and consequently the blast locations are not well-constrained. This can be observed as broader amplitude maxima to the West and East of Oslo, as well as biases with respect to the ground-truth locations. Note that since OSLN4 was not deployed before October 2022 and produced corrupted data from summer 2023 onwards, poor resolution is also expected for locations at the Huseby construction site for events outside this time interval. As a consequence, blasts from the Fornebu metro tunnel and the water storage site cannot always be discriminated. However, as the comparison with ground-truth location shows, the accuracy is very good when the entire network was in operation (see Fig. 5, 6 and 7).

Figure 8a shows all locatable outliers triggered at OSLN2 and EKBG1 in the study period (1,272 events). Two different symbol colors are used to distinguish the time period of complete and incomplete station coverage (see Table 1). Almost all events are located inside or close to areas of known construction or quarry blast activities. Clusters of events are indicated, of which we have already seen examples above. By far the most blasting activity is observed to the West of the city center, i.e., the Huseby and Fornebu constructions sites. Events are well-located during complete station coverage, while a number of blasts tend to be falsely located westwards from Huseby when the westernmost station OSNL4 was not in operation. Note that we have filtered out distant events (blasts from distant quarries and regional and teleseismic earthquakes) since they are usually falsely located at the edge of the grid search region or in the center of the network if the incidence angles are steeper.

Figure 8b shows the corresponding time line of locatable outliers. Before July 2022 only stations with high noise levels in the east of Oslo were in operation and consequently only a few events are observed. For the rest of the study period, there is a lot of blasting activity with up to 10 events per day and about 4 per day on average. Pauses in the blasting during public holidays (Easter and Christmas break) and school holidays during summer are clearly visible. Figure 9 presents more detection statistics for all located blasts. Time of day and day of week distribution are consistent with blasting which usually ceases on Sundays and during nighttime (Figure 9d and e). Local seismic magnitudes of blasts are between -0.5 and 1.5 (Figure 9c).

The temporal distribution of blasts in the reviewed reference data set is shown in the background of Figure 8b. A high percentage of these events are recognized as outliers (69%). However, 31% of visually identified blasts are not found (i.e., false negatives; 520 events). A closer look at the distributions of STA/LTA ratios of all events in Figure 9a and b, as a proxy for SNRs, gives an explanation for those results. The distribution of ratios are shown for all locatable outlier events, for all detected outliers (including those that were not locatable), and for all STA/LTA detections. Note that logarithmic scales for number of detections are used. The comparison shows that STA/LTA detections at OSLN2 with STA/LTA ratios above 10 are almost all classified as outliers, the majority being also locatable as blasts (Fig. 8a). In other words, there would be no need for an outlier detection method for those events, and we could simply use the STA/LTA detections directly to monitor blasts. However, towards lower SNRs the picture changes. We obtain an increasing number of STA/LTA detections, the

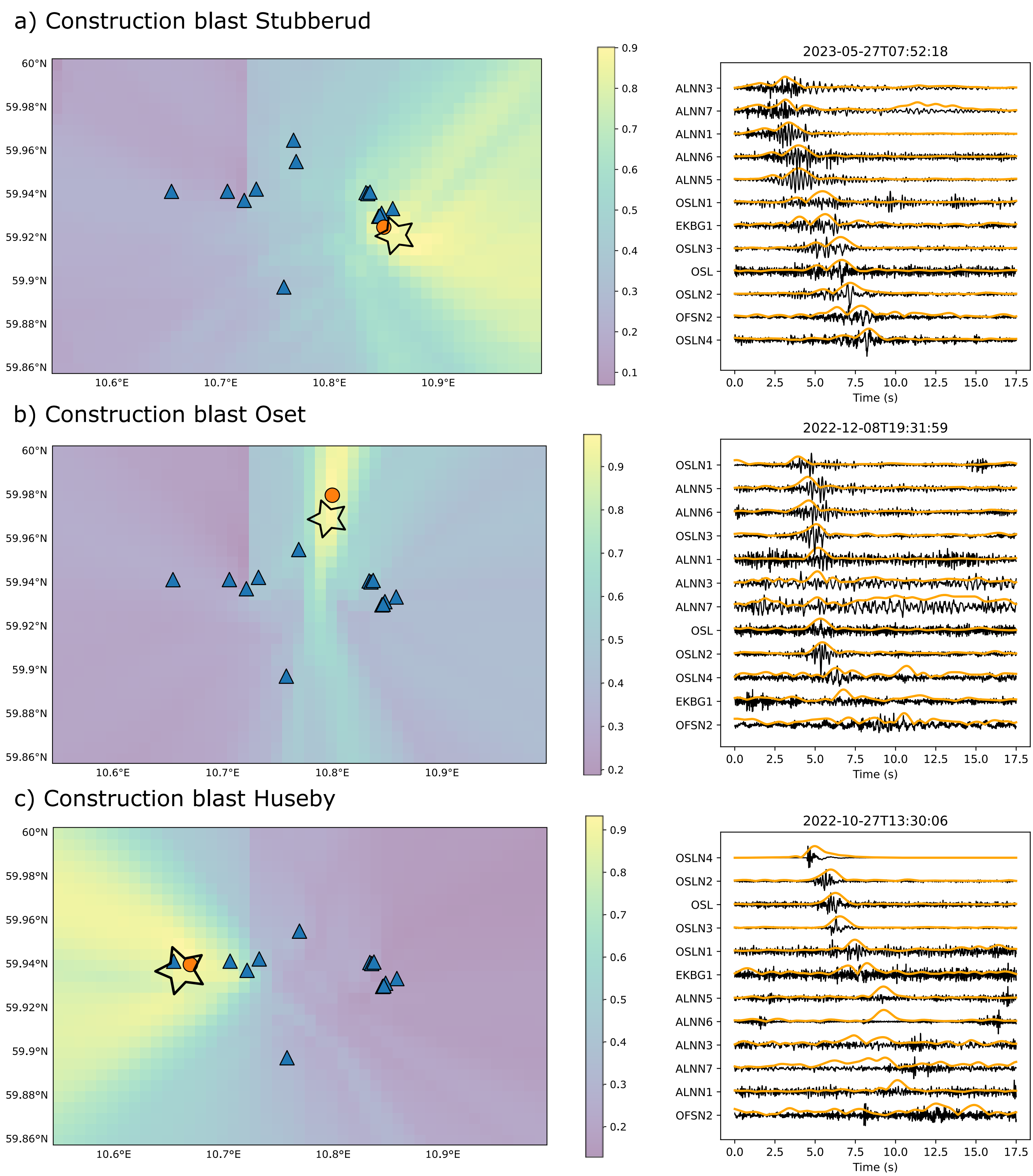


Figure 5 Map of Oslo region overlayed with stacked envelope amplitudes (normalized). High amplitudes indicate more likely event location. The best location (orange cricle) and seismic stations (triangles) are shown. The star symbols and their extents indicate the areas of known blasting activity. On the right-hand side the vertical component seismic data for all stations are shown. Orange data indicate low-pass filtered envelopes enhancing Rg arrivals.

majority not being locatable events, probably mostly local noise bursts. The outlier detection method allows us to reduce the number of detections to be screened for location considerably. For the lowest SNRs we obtain a number of locatable STA/LTA detections which turned out to be blasts not detected as outliers (red bars). This indicates that our outlier detection method has limitations in recognizing weak events. We will deal with these missed events when applying the supervised blast classifier. It is worth emphasizing that visual inspection of the unlocatable outliers at OSLN2 revealed clear blast signals which were not observed on more than three stations. Four examples are included in supplementary Figure S6. This shows that the outlier detector com-

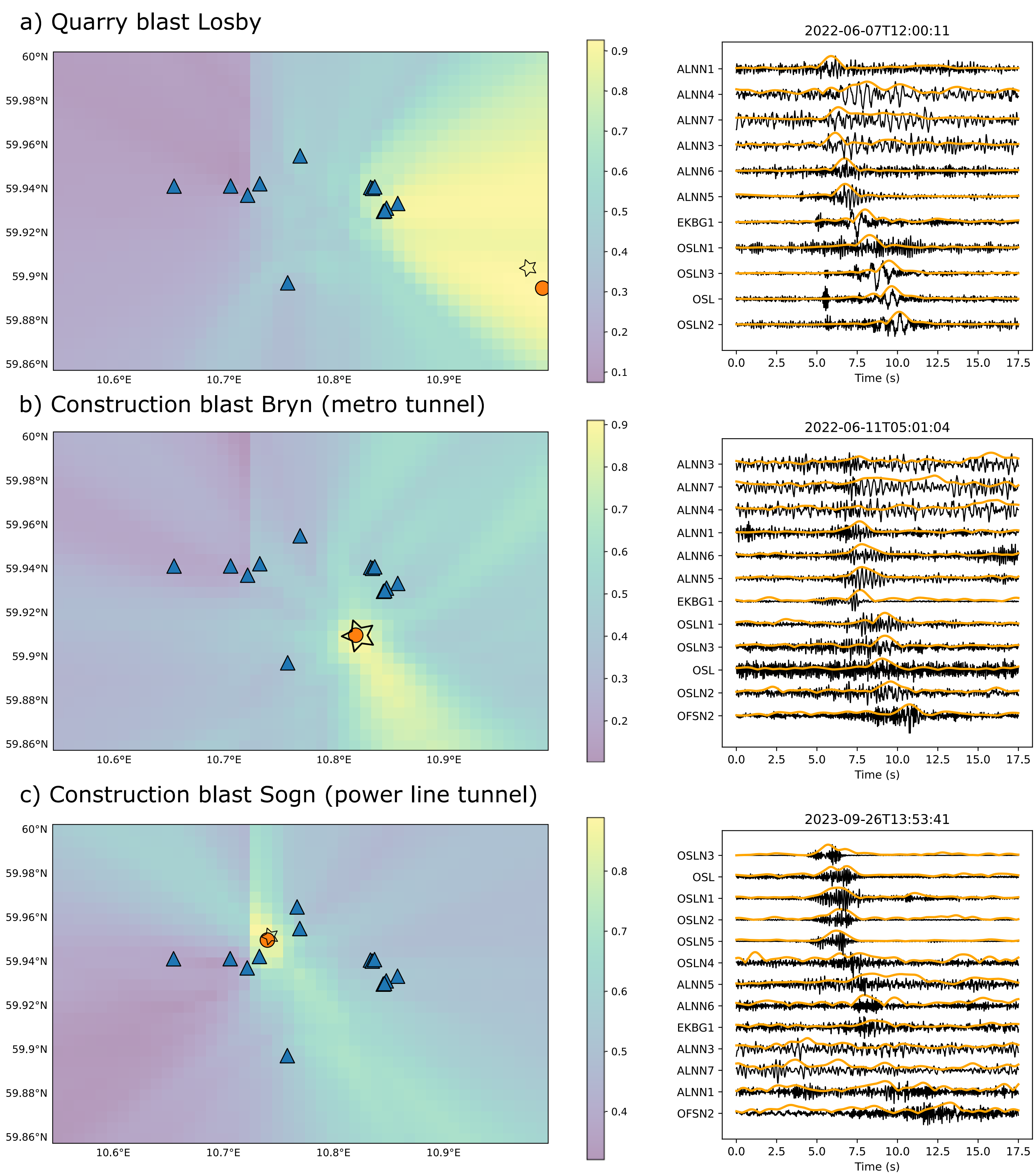


Figure 6 Same as Figure 5. Note that smaller size of star in (c) indicates more certain (ground-truth) blast location.

bined with a denser station network would have recognized even more blasts.

For EKBG1 (Fig. 8b) there are a few more detections with high SNRs that are not classified as outliers. The major difference to OSLN2 is that fewer outliers turned out to be locatable events which could be identified as blasts. In other words, a lot of outliers are seismic events at EKBG1 which are not observed on other stations. Given that we know that blasts are usually picked up by at least two more stations, it is likely that these are mostly local noise bursts around EKBG1 that cannot be explained by normal background noise fluctuations and hence are not well-reconstructed by the auto-encoder. However, it is worth emphasizing that many STA/LTA are still removed by outlier detection, which reduces the amount of detections to be screened for possible localization.

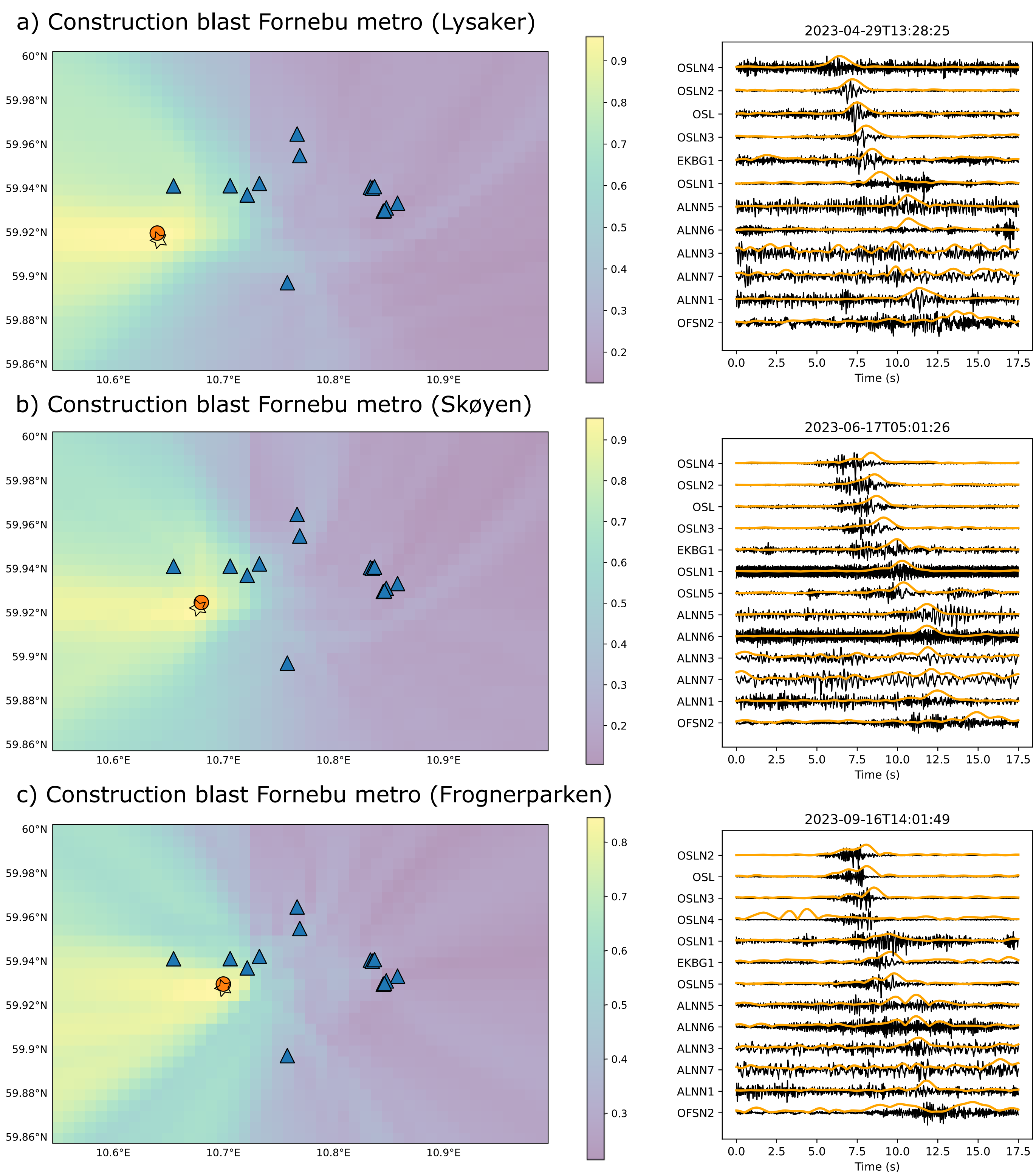


Figure 7 Same as Figure 5. Note that smaller size of stars in all panels indicate more certain (ground-truth) blast locations.

5 Results of blast classifier

Ideally, we need to train a classifier with data from several stations so that it generalizes well enough and can identify blast signals in the data from different stations. As also discussed below, the reason for starting with a classifier trained for a single station (OSLN2) is the limited number of blasts observed on all stations which results in an unbalanced training data set with respect to event location and observing stations. However, a few stations close to OSLN2 have a comparable number of observations and could be included in the training. We will come back to this in the discussion section.

We train the blast classifier with waveforms from OSLN2 including all 1,272 blasts from different areas in Oslo detected as outliers at OSLN2 and with the same number of noise examples. The reason for not using all 1,870 signals in the reference data set is that we want to simulate a workflow where we only train with blasts previously detected by the outlier detector, excluding those

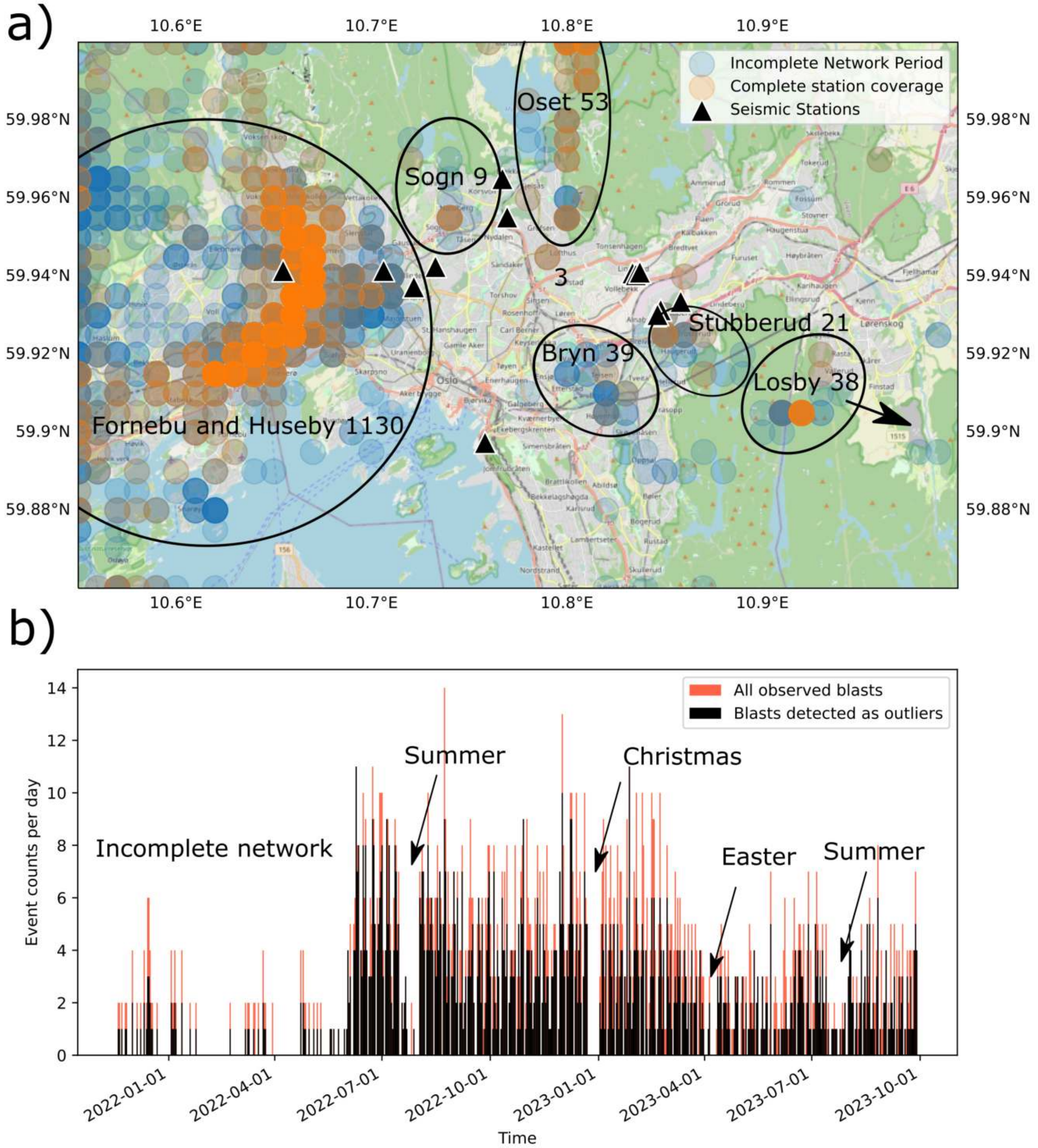


Figure 8 a) Located seismic events from outlier detection. Clusters and number of blasts are indicated. b) Time line of those events together with all observed and manually confirmed blast STA/LTA detections.

that we only identified after screening all STA/LTA detections. The classification performance metrics of the best model of all five folds using data not used for training are shown in Table 2. We achieve high values for both precision and recall.

Class	Blast	Not blast
Precision	0.92	0.95
Recall	0.94	0.93
F1	0.93	0.94
Accuracy	0.93	

Table 2 Performance of blast classifier on validation data.

Next, we apply the classifier to all 29,058 STA/LTA detections at OSLN2 in the time period from 01/06/2022 until 28/09/2023. We use a probability threshold of 0.5, i.e., picking the winning class (blast vs. not blast) as well as 0.7 to test different confidence levels. Figure 10a shows the time line of 1,385 classified and locatable blasts using a threshold of 0.7 together with the refer-

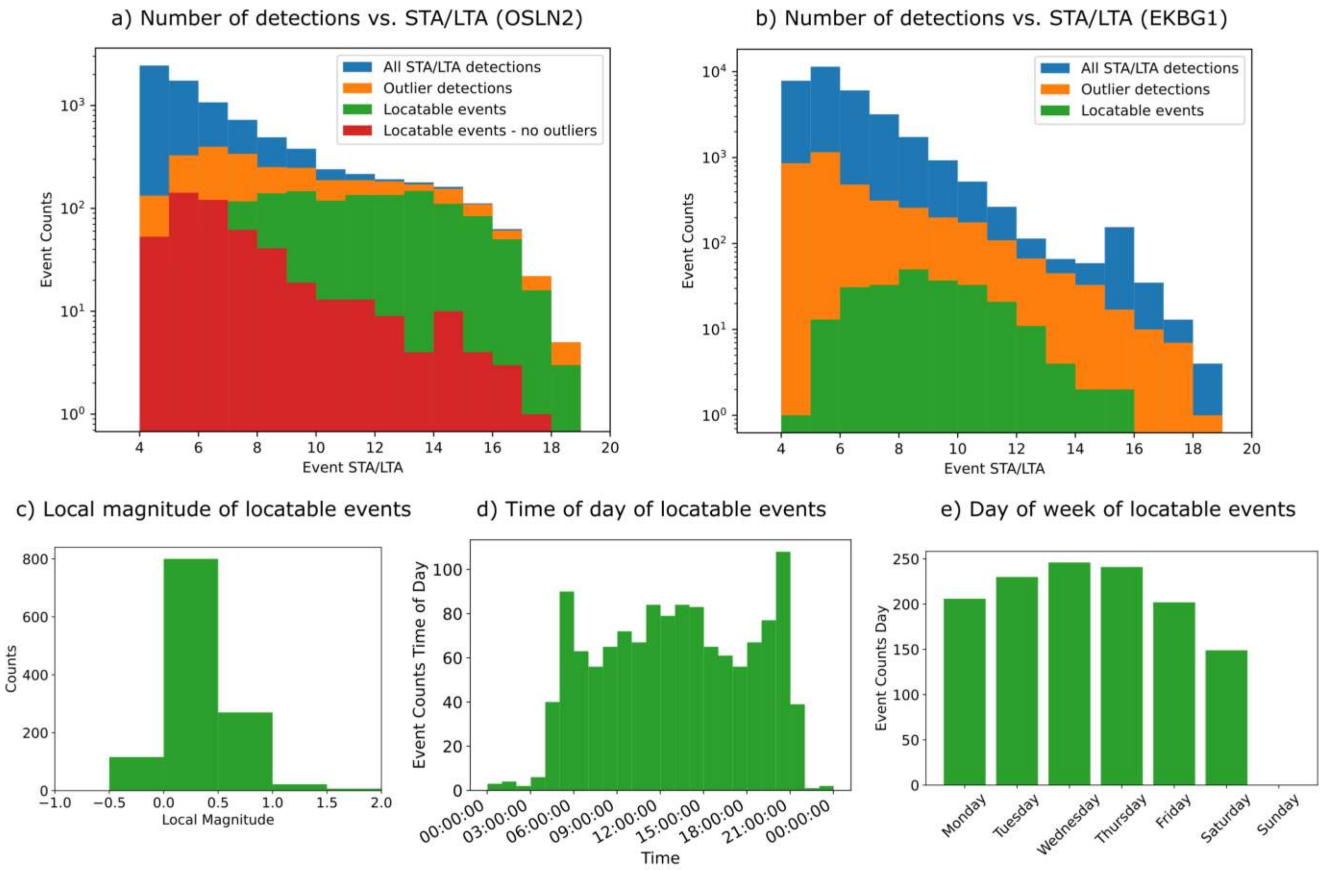


Figure 9 Statistics of STA/LTA and outliers detections, and locatable blasts.

ence data set. In comparison with the outlier detector (Fig. 8b), more blasts are recognized. When we use our reference data set, which includes 1,870 blast manually identified by screening all locatable STA/LTA detections, as ground-truth for evaluation, the recognition rate increases from 69 to 80%, and the number of False Negatives decreases to 371 events. A number of 22 False Positives are either blasts outside the study area, which we did not include in the reference data set of locatable blasts, or are local signals at OSLN2 which are randomly associated with Rg wave-like signals at other stations. Of all STA/LTA detections, 424 events are classified as blasts but are not locatable. As with the outlier detector, these events are not necessarily false, but are simply not observed on more than three stations, which would allow for a reliable location and for being included in the reference data set (supplementary Figure S6 shows examples). In fact, we checked a selection of these detections manually and found that almost all of these show clear blast-like signatures at station OSLN2 and OSLN3. Hence, in relation to the high number of tested STA/LTA detections (29,058), the actual number of false classification is negligible if the goal is to provide real-time information about ongoing blasting at construction sites. However, if the goal is early warning in case of unusual events (accidents, attacks), any false detection should be avoided and other data have to be included before issuing an alert.

For a probability threshold of 0.5 the blast recognition rate increases further to 87%. As the maps in Figure 10b–c show, this is partly due to more of the underrepresented blasts in the north and east of Oslo being correctly classified (compare Fig. 10b and c). It is expected that those events yield a lower blast probability since the training data is unbalanced with respect to event location. However, there are also more False Positives (77) and about 100 more unlocatable events (519) compared to using the high threshold. Visually inspecting those 100 additional events revealed that about 30% look like blasts, but are not locatable because they are observed on only one or two stations. However, the rest (70%) are now actual false detections which we would avoid with a higher threshold.

6 Discussion

We present a prototype for an automatic urban seismic monitoring system which identifies any potentially interesting event as well as routinely detects previously identified events. Our system is based on a low number of low-cost seismic sensors and was running for almost two years. We demonstrate that with comparably low effort when it comes to upgrades of the sensor infrastructure, a city can be monitored continuously for explosion activity. In our case this includes construction blasts, but there is no reason for other types of explosions, such as accidents or deliberate attacks, not to be detected as long as there is sufficient coupling with the ground.

An alternative approach to identify events of inter-

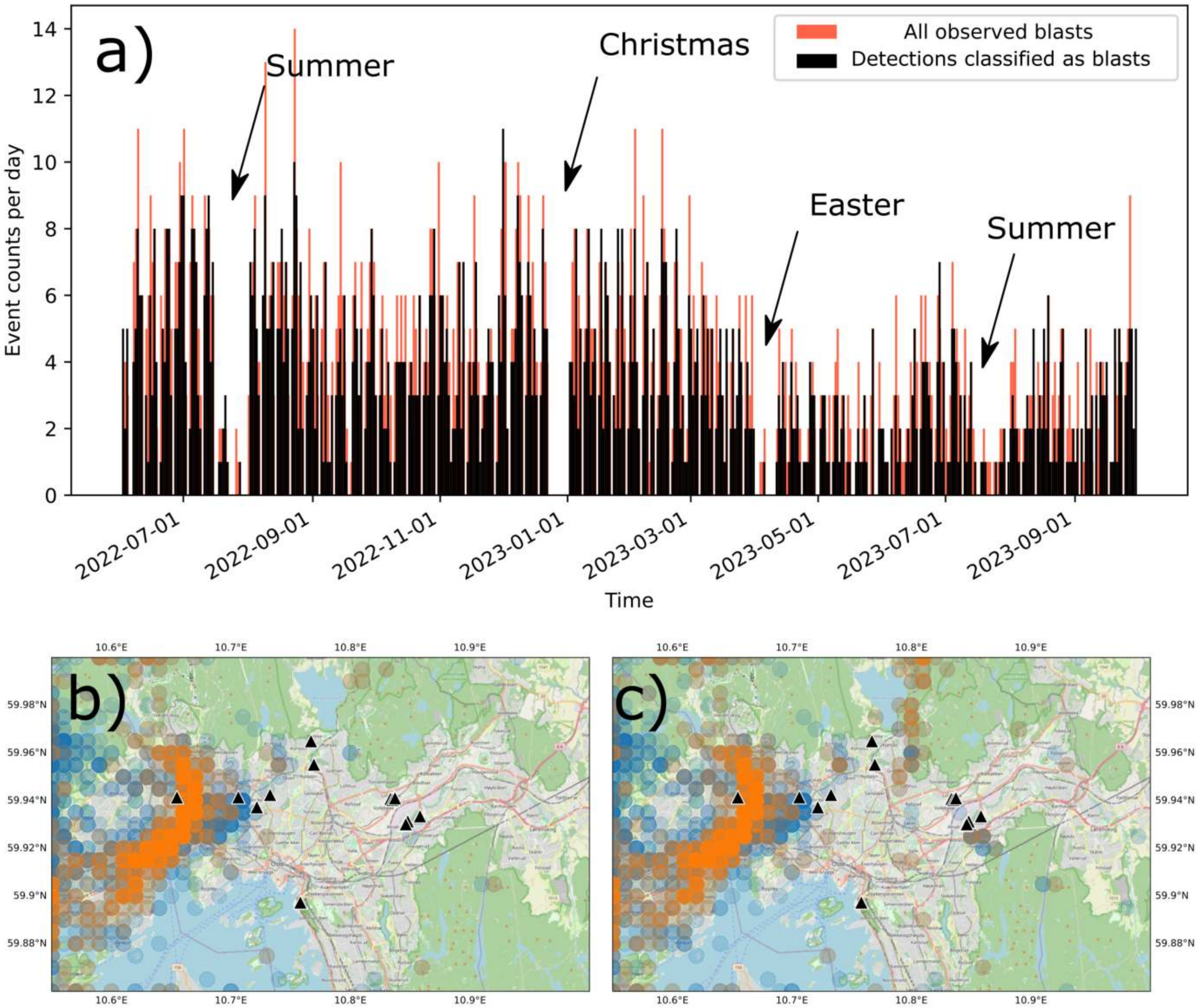


Figure 10 a) Time line of locatable STA/LTA detections classified as blasts by the CNN model in comparison to reference blast data set. b–c) Locations of classified blasts using a probability threshold of 0.7 (a) and 0.5 (c). Orange symbols are for complete and blue symbols for incomplete network operation. Black triangles show seismic stations.

est in an unsupervised fashion would be to apply recently developed (deep) clustering methods (e.g., Seydoux et al., 2020). While adapting such methods and comparison of the results with our method would go beyond the scope of the current paper, we encourage further studies to compare both approaches. The main reason why we did not choose a method to identify outlier events via clustering, is the additional need to identify the cluster(s) of interest and the risk that outlier might be too rare to form separate clusters. We found that an auto-encoder is a relatively easily implemented and trained alternative (compared to more sophisticated deep learning models), and is not difficult to tune for each station. The only tuning we did for each new station was the selection of the training data, the number of samples in the input time window, and the latent dimension if the auto-encoder. We identified the sample number and latent dimension to be the most important parameters, while the rest of the model architecture, hyper-parameters, and the choice of training data (except that it should cover different noise conditions at different times) does not need to be adapted for each station.

There are different possibilities for improving the system. First and most important, the seismic station network can be extended by covering a larger area and by increasing station density. This will improve location accuracy, especially outside the current network area beyond the city limits. A denser sensor deployment will also enable locating more events that so far are only observed on a single or two stations and are, therefore, not locatable with our Rg wave stacking approach. This would also allow us to potentially run the outlier detector on more than two stations simultaneously and to locate the detected events with more stations that are close-by.

Secondly, the detection process could be further improved. The outlier detector could be retrained regularly since local noise conditions around each station may have changed over time. Furthermore, more work can be done to further tune the neural network architecture of the auto-encoder to optimize outlier detection at different seismic stations. One could also further investigate if the same outlier detector model can be applied to different stations despite of station-specific background noise conditions and sub-surface-related site effects. To test this we already trained an additional auto-encoder model with combined data from stations OSLN2 and OSLN3. We then applied this and the original model for OSLN2 to both stations and compared the results. The outlier detector performance confirmed our choice to make the detector station-specific (see recall-precision curves in supplementary Figure S2). While we found best performance with the current model, we acknowledge the high number of latent features. However, this is not an issue as such since our goal is not data compression but outlier detection. A more systematic study of outlier detector performance for decreasing the number of latent features beyond the tests we presented here may help to further optimize the detection rate.

We train the supervised blast classifier with a comparably low number of data points. Longer recording periods will increase the available data and, thus, most likely improve classifier performance. Alternatively, it is possible to augment the existing training data with noise or other seismic signals (Köhler et al., 2022). This is of particular importance for areas with infrequent blasts which are currently not well-represented in the training data and are therefore less likely to be correctly classified (events in the east of Oslo). We only used a single station to train the classifier. Consequently, the model learns station-specific features and does not generalize well. However, ideally we would need a classifier which generalizes well enough to detect blasts on all stations. We started to explore training a single classifier with waveform data from all stations, either using independent input data (three-component waveforms from different stations) or using multiple channels from different stations in one data sample as input. However, we found that both approaches require more and better balanced training data when it comes to blast detection. We trained a model using data from two stations in the west of Oslo (OSLN2 and OSLN3). Supplementary Figure S3 shows that the generalization ability is not satisfactory. The classifier trained on OSLN2 and applied to OSLN3 does not perform well at all. The model trained on both stations and applied separately to both stations performs better at OSLN3, but still clearly worse than our preferred model. The model trained on both stations and applied to OSLN2 performs slightly worse than the original model we used above. With longer time series of blasts being available in future, we would like to generalize the blast classifier for more stations, and most important, for other source areas.

The SNRs of many detected blasts in our study are rather low (see Fig. 9a and b) which is expected for an urban environment. This naturally impacts the performance of the outlier detector, as discussed above, as well as to some extent the blast classifier. Again, the best way to deal with this issue is a larger training data set to better represent noisy waveform data.

For evaluating our methodology we used a manually compiled data set of locatable blasts in the city of Oslo. We have shown that the outlier detector detects about 70% of those events. The classifier specifically trained to detect blasts increases recognition rate to 80%. If the goal of monitoring is to detect as many real blasts as possible while accepting a number of false alarms, the classifier performance can be improved further up to 87% by using a lower probability threshold. In general, our outlier event detector and blast classifier generate a very low number of false alarms. We encountered randomly associated Rg waves producing false events in less than 1% of the locatable outlier events. However, this is partially due to the combination with the Rg-based locations procedure which sorts out many unlocatable events. Nevertheless, even without event location, the amount of events to be processed is reduced considerably compared to simply applying an STA/LTA detector and attempting to locate all those events. Furthermore, we found also many real blast signals among the outliers being unlocatable due to limited station density. With larger training data and denser seismic networks, we therefore expect the benefit of our methodology to become even more evident.

A system such the one we have proposed has to be adapted and modified when deployed in another city. The most critical parts are the existence of a station network with sufficient resolution for event location and the deployment of a functional outlier detector. If all signals of interest have a high SNR, it may be sufficient to simply use an STA/LTA detector combined with Rg-wave based location for outlier detection. However, the supervised blast classifier would still be part of such a system. Before deploying the network and selecting stations for the outlier detector, potential source locations for blasts, e.g., construction sites, or infrastructure to be monitored should be identified. From our experience, the Rg waves of blasts needed for localization can be observed at up to about 12 km distance, whereas blast signals detected as outliers require stations at not more than about 6 km distance. Furthermore, adaption to other stations in another city requires retraining of the DL models. The outlier detector does not require a large data set for this. After 2-3 weeks there should be sufficient wave field variability captured to train the auto-encoder(s). However, the blast classifier would require rather frequent blasting to gather enough events for training the CNN; from our experience around 1,000 events are needed. This might be a limitation of our work flow in areas with infrequent blasting.

The final question is how our system can be integrated into a smart-city solution. The simplest objective could be to provide the general public with real-time information about event locations on a public web-dashboard or through a mobile app. If citizens felt ground shaking, they can easily check if it was related to any known construction site. If the goal is early warning in case of unusual events (accidents, attacks) and the seismic monitoring system is supposed to automatically alert the city authorities and possibly the public, combination with other data sources available to the stakeholders may be needed to avoid false alarms. All data layers combined could then provide automatic alerts and initiate further actions. For example, we envision that the recorded ground motion could be utilized to predict potential damage of infrastructure and buildings categorized on the European Macroseismic Scale (EMS-98). If most detected events are construction blasts as in our case, automatic monitoring can still be useful in a smart-city application, for example to alert about blasts above a certain magnitude or amplitude threshold in areas with unstable ground such as quick clay.

7 Conclusions

The objective of this study was to detect events in the city of Oslo, Norway, that generate seismic signals. To this end, we have successfully developed a prototype of an automatic urban seismic monitoring system using input data from low-cost seismic sensors deployed between 2021 and 2023. The work flow of our system includes two deep learning methods: the first one identifies rare events using an event outlier detector based on an auto-encoder model and the second one classifies events of interest using a CNN model trained in a supervised manner. Both methods used waveforms of signals as input, which were pre-detected using the traditional STA/LTA trigger method. For both evaluating the outlier detector and training the event classifier with events of interest, we relied on locating the seismic signals using Rg waves observed on the seismic network.

The results of our approach impressively reveal ongoing construction activity and their temporal variation in the city of Oslo. From about 1,870 construction blasts in different areas during 22 months of monitoring generating locatable seismic signals on our network, the outlier detector recognized 69%. The classifier trained on these blasts was able to detect between 80 and 87% of those, many with low Signal-to-Noise ratios. At the same time, the false detection rate is very low. In absolute numbers, the automatic system was able to retrieve 1,271 of the manually identified blasts in the initial outlier detection step, and between 1,385 and 1,627 blasts, depending on the detection threshold, using the blast classifier.

The performance of our prototype system could be improved by expanding and densifying the seismic network as well as increasing the training data with more blast records. However, we demonstrated that even with a low number of seismic sensors, a city can be monitored automatically and continuously for explosion events. This opens up new possibilities to include seismic records into the sensor data stream of future smart city solutions. We are therefore confident that the outcome of our pilot study represents a robust prototype system for urban explosion monitoring in the city of Oslo and possibly elsewhere.

Acknowledgements

Seismic data processing was done using Obspy (Beyreuther et al., 2010) and DL models were implemented in Keras (Chollet et al., 2015). This study was funded by The Research Council of Norway (project number 311596, GEObyIT). We thank all seismic station hosts for participating in this study: Oslo municipality, Møller Medvind AS, Tore Ligård AS, Gravco AS, Alfaset graveyard, Bjerke VGS, Linderud skole, Kongshavn VGS, Ullevål skole, Vinderen skole, Hovseter skole, and the Gleiss family. Special thanks go to the NORSAR technicians who helped deploying and maintaining the network, in particular Sindre Stokkan and Bjørn Christian Freitag. We thank Karoline Aastrup-Köhler for spellchecking the manuscript. We thank René Steinmann and one anonymous reviewer for their very helpful comments and suggestions.

Data and code availability

All data of the temporary network and station OSL are openly available at the Norwegian EIDA node (Ottemöller et al., 2021). The temporary stations have the network code 4X (Köhler, 2021). Information about blasting activity in Oslo was retrieved from https://nabovarsling.no. Maintained code is available at https://github.com/NorwegianSeismicArray/urban-seismic-monitoring. The version of the code at the time of publication of this study with scripts accessing the data on the Norwegian EIDA node is available at https://doi.org/10.5281/zenodo.10777734 (Köhler and Myklebust, 2024).

Competing interests

The authors have no competing interests.

References

Al Nuaimi, E., Al Neyadi, H., Mohamed, N., and Al-Jaroodi, J. Applications of big data to smart cities. *Journal of Internet Services and Applications*, 6(1):1–15, 2015. doi: 10.1186/s13174-015-0041-5.

Bergen, K. J., Chen, T., and Li, Z. Preface to the Focus Section on Machine Learning in Seismology. *Seismological Research Letters*, 90(2A):477–480, 2019. doi: 10.1785/0220190018.

Bergen University, E. S. Operation of the Norwegian National Seismic Network 2011. Technical report, Department of Earth Science University of Bergen, 2012. https://nnsn.geo.uib.no/reports/2011/All_reports_2011.pdf.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., and Wassermann, J. ObsPy: A Python toolbox for seismology. *Seismological Research Letters*, 81(3):530–533, 2010. doi: 10.1785/gssrl.81.3.530.

Bouchard, S., L'Heureux, J.-S., Johansson, J., Leroueil, S., and LeBoeuf, D. Blasting induced landslides in sensitive clays. In *Landslides and Engineered Slopes. Experience, Theory and Practice*, pages 497–504. CRC Press, 2018.

Brockerhoff, M. Urban growth in developing countries: a review of projections and predictions. *Population and development Review*, 25(4):757–778, 1999. doi: 10.1111/j.1728-4457.1999.00757.x.

Chamarczuk, M., Nishitsuji, Y., Malinowski, M., and Draganov, D. Unsupervised learning used in automatic detection and classification of ambient-noise recordings from a large-N array. *Seismological Research Letters*, 91(1):370–389, 2020. doi: 10.1785/0220190063.

Chang, S. E., McDaniels, T., Fox, J., Dhariwal, R., and Longstaff, H. Toward disaster-resilient cities: Characterizing resilience of infrastructure systems with expert judgments. *Risk analysis*, 34 (3):416–434, 2014. doi: 10.1111/risa.12133.

Chollet, F. et al. Keras. https://github.com/fchollet/keras.

Dando, B. D., Goertz-Allmann, B. P., Brissaud, Q., Köhler, A., Schweitzer, J., Kværna, T., and Liashchuk, A. Identifying attacks in the Russia–Ukraine conflict using seismic array data. *Nature*, pages 1–6, 2023. doi: 10.1038/s41586-023-06416-7.

Díaz, J., Ruiz, M., Sánchez-Pastor, P. S., and Romero, P. Urban seismology: On the origin of earth vibrations within a city. *Scientific reports*, 7(1):15296, 2017. doi: 10.1038/s41598-017-15499-y.

Dowding, C. H. *Blast Vibration Monitoring for Engineering*, volume 4. Elsevier, 2016.

Fiori, R., Vergne, J., Schmittbuhl, J., and Zigone, D. Monitoring induced microseismicity in an urban context using very small seismic arrays: The case study of the Vendenheim EGS project. *Geophysics*, 88(5):WB71–WB87, 2023. doi: 10.1190/geo2022-0620.1.

Fischer, J., Redlich, J.-P., Scheuermann, B., Schiller, J., Günes, M., Nagel, K., Wagner, P., Scheidgen, M., Zubow, A., Eveslage, I., et al. From earthquake detection to traffic surveillance–about information and communication infrastructures for smart cities. In *System Analysis and Modeling: Theory and Practice: 7th International Workshop, SAM 2012, Innsbruck, Austria, October 1-2, 2012. Revised Selected Papers 7*, pages 121–141. Springer, 2013. doi: 10.1007/978-3-642-36757-1_8.

Gharti, H. N., Oye, V., Roth, M., and Kühn, D. Automated microearthquake location using envelope stacking and robust global optimization. *Geophysics*, 75(4):MA27–MA46, 2010. doi: 10.1190/1.3432784.

Gibbons, S. J. and Ringdal, F. The detection of low magnitude seismic events using array-based waveform correlation. *Geophysical Journal International*, 165(1):149–166, 2006. doi: 10.1111/j.1365-246X.2006.02865.x.

Hillers, G., T. Vuorinen, T. A., Uski, M. R., Kortström, J. T., Mäntyniemi, P. B., Tiira, T., Malin, P. E., and Saarno, T. The 2018 geothermal reservoir stimulation in Espoo/Helsinki, southern Finland: Seismic network anatomy and data features. *Seismological Research Letters*, 91(2A):770–786, 2020. doi: 10.1785/0220190253.

Jenkins, W. F., Gerstoft, P., Bianco, M. J., and Bromirski, P. D. Unsupervised deep clustering of seismic data: Monitoring the Ross Ice Shelf, Antarctica. *Journal of Geophysical Research: Solid Earth*, 126(9):e2021JB021716, 2021. doi: 10.1029/2021JB021716.

Johnson, C. W., Ben-Zion, Y., Meng, H., and Vernon, F. Identifying different classes of seismic noise signals using unsupervised learning. *Geophysical Research Letters*, 47(15):e2020GL088353, 2020. doi: 10.1029/2020GL088353.

Kalinowski, M. B. and Mialle, P. Introduction to the topical issue on nuclear explosion monitoring and verification: scientific and technological advances. *Pure and Applied Geophysics*, 178(7): 2397–2401, 2021. doi: 10.1007/s00024-021-02783-2.

Köhler, A., Ohrnberger, M., and Scherbaum, F. Unsupervised pattern recognition in continuous seismic wavefield records using self-organizing maps. *Geophysical Journal International*, 182(3): 1619–1630, 2010. doi: 10.1111/j.1365-246X.2010.04709.x.

Köhler, A., Myklebust, E., and Mæland, S. Enhancing seismic calving event identification in Svalbard through empirical matched field processing and machine learning. *Geophysical Journal International*, 230(2):1305–1317, 2022. doi: 10.1093/gji/ggac117.

Kong, Q., Allen, R. M., Schreier, L., and Kwon, Y.-W. MyShake: A smartphone seismic network for earthquake early warning and beyond. *Science advances*, 2(2):e1501055, 2016. doi: 10.1126/sciadv.1501055.

Kong, Q., Trugman, D. T., Ross, Z. E., Bianco, M. J., Meade, B. J., and Gerstoft, P. Machine learning in seismology: Turning data into insights. *Seismological Research Letters*, 90(1):3–14, 2019. doi: 10.1785/0220180259.

Kraft, T., Mai, P. M., Wiemer, S., Deichmann, N., Ripperger, J., Kästli, P., Bachmann, C., Fäh, D., Wössner, J., and Giardini, D. Enhanced geothermal systems: Mitigating risk in urban areas. *Eos, Transactions American Geophysical Union*, 90(32):273–274, 2009. doi: 10.1029/2009EO320001.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. ImageNet classification with deep convolutional neural networks. In Pereira, F., Burges, C., Bottou, L., and Weinberger, K., editors, *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc., 2012. https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.

Köhler, A. GEObyIT seismic network in Oslo, Norway, 2021. doi: 10.7914/3dms-sj84.

Köhler, A. and Myklebust, E. B. Code for monitoring urban construction and quarry blasts with low-cost seismic sensors and machine learning tools in the city of Oslo, Norway, 2024. doi: 10.5281/zenodo.10777734.

Leonard, M. and Kennett, B. Multi-component autoregressive techniques for the analysis of seismograms. *Physics of the Earth and Planetary Interiors*, 113(1-4):247–263, 1999. doi: https://doi.org/10.1016/S0031-9201(99)00054-0.

McKinsey. *Smart cities: Digital solution for a more liveable future*. NY: McKinsey Global Institute, 2018.

Mousavi, S. M. and Beroza, G. C. Machine Learning in Earthquake Seismology. *Annual Review of Earth and Planetary Sciences*, 51, 2023. doi: 10.1146/annurev-earth-071822-100323.

Mousavi, S. M., Zhu, W., Ellsworth, W., and Beroza, G. Unsupervised clustering of seismic signals using deep convolutional autoencoders. *IEEE Geoscience and Remote Sensing Letters*, 16(11): 1693–1697, 2019. doi: 10.1109/LGRS.2019.2909218.

Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., and Beroza, G. C. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature communications*, 11(1):1–12, 2020. doi: 10.1038/s41467-020-17591-w.

Navarro, J., Schiavon, A., Vieira, M., and Silva, P. Deep seismic compression. In *81st EAGE Conference and Exhibition 2019*, pages 1–5. European Association of Geoscientists & Engineers, 2019. doi: 10.3997/2214-4609.201901620.

Naveen, G., Sastry, V., and Chandar, K. R. Assessment of Structural Damage Due to Blasting in Hydro Power Tunnel. In *International Conference on Geotechnical Challenges in Mining, Tunneling and Underground Infrastructures*, pages 229–240. Springer, 2021.

Nugent, J. Raspberry Shake: Watch the Earth move under your feet. *Science Scope*, 42(4):22–25, 2018. https://www.jstor.org/stable/26611884.

Nuha, H. H., Balghonaim, A., Liu, B., Mohandes, M., Deriche, M., and Fekri, F. Deep neural networks with extreme learning machine for seismic data compression. *Arabian Journal for Science and Engineering*, 45(3):1367–1377, 2020. doi: 10.1007/s13369-019-03942-3.

OpenStreetMap contributors. Planet dump retrieved from https://planet.osm.org , 2017. https:

//www.openstreetmap.org.

Ottemöller, L., Strømme, M. L., and Storheim, B. M. Seismic monitoring and data processing at the Norwegian National Seismic Network. *Summary of the Bulletin of the International Seismological Centre*, 52(I):27–40, 2018.

Ottemöller, L., Michalek, J., Christensen, J.-M., Baadshaug, U., Halpaap, F., Natvik, Ø., Kværna, T., and Oye, V. UiB-NORSAR EIDA node: Integration of seismological data in Norway. *Seismological Society of America*, 92(3):1491–1500, 2021. doi: 10.1785/0220200369.

Provost, F., Hibert, C., and Malet, J.-P. Automatic classification of endogenous landslide seismicity using the Random Forest supervised classifier. *Geophysical Research Letters*, 44(1):113–120, 2017. doi: 10.1002/2016GL070709.

Ritter, J. R., Balan, S. F., Bonjer, K.-P., Diehl, T., Forbriger, T., Mărmureanu, G., Wenzel, F., and Wirth, W. Broadband urban seismology in the Bucharest metropolitan area. *Seismological Research Letters*, 76(5):574–580, 2005. doi: 10.1785/gssrl.76.5.574.

Seydoux, L., Balestriero, R., Poli, P., Hoop, M. d., Campillo, M., and Baraniuk, R. Clustering earthquake signals and background noises in continuous seismic data with unsupervised deep learning. *Nature communications*, 11(1):3972, 2020. doi: 10.1038/s41467-020-17841-x.

Shallan, O., Eraky, A., Sakr, T., and Emad, S. Response of building structures to blast effects. *International journal of engineering and innovative technology*, 4(2):167–175, 2014. https://www.humanitarianlibrary.org/sites/default/files/2018/10/IJEIT1412201408_30.pdf.

Sick, B., Guggenmos, M., and Joswig, M. Chances and limits of single-station seismic event clustering by unsupervised pattern recognition. *Geophysical Journal International*, 201(3):1801–1813, 2015. doi: 10.1093/gji/ggv126.

Spica, Z. J., Perton, M., Martin, E. R., Beroza, G. C., and Biondi, B. Urban seismic site characterization by fiber-optic seismology. *Journal of Geophysical Research: Solid Earth*, 125(3):e2019JB018656, 2020. doi: 10.1029/2019JB018656.

Steinmann, R., Seydoux, L., Beaucé, E., and Campillo, M. Hierarchical exploration of continuous seismograms with unsupervised learning. *Journal of Geophysical Research: Solid Earth*, 127(1):e2021JB022455, 2022a. doi: 10.1029/2021JB022455.

Steinmann, R., Seydoux, L., and Campillo, M. AI-based unmixing of medium and source signatures from seismograms: ground freezing patterns. *Geophysical Research Letters*, 49(15):e2022GL098854, 2022b. doi: 10.1029/2022GL098854.

Thill, M., Konen, W., Wang, H., and Bäck, T. Temporal convolutional autoencoder for unsupervised anomaly detection in time series. *Applied Soft Computing*, 112:107751, 2021. doi: 10.1016/j.asoc.2021.107751.

Valentine, A. P. and Trampert, J. Data space reduction, quality assessment and searching of seismograms: autoencoder networks for waveform data. *Geophysical Journal International*, 189(2):1183–1202, 2012. doi: 10.1111/j.1365-246X.2012.05429.x.

Wang, Y., Yao, H., and Zhao, S. Auto-encoder based dimensionality reduction. *Neurocomputing*, 184:232–242, 2016. doi: 10.1016/j.neucom.2015.08.104.

Yin, C., Zhang, S., Wang, J., and Xiong, N. N. Anomaly detection based on convolutional recurrent autoencoder for IoT time series. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 52(1):112–122, 2020. doi: 10.6038/pg2020DD0013.

Yoon, C. E., O'Reilly, O., Bergen, K. J., and Beroza, G. C. Earthquake detection through computationally efficient similarity search. *Science advances*, 1(11):e1501057, 2015. doi: 10.1126/sciadv.1501057.

Zhang, K., Ni, J., Yang, K., Liang, X., Ren, J., and Shen, X. S. Security and privacy in smart city applications: Challenges and solutions. *IEEE Communications Magazine*, 55(1):122–129, 2017. doi: 10.1109/MCOM.2017.1600267CM.

Zheng, H. and Zhang, B. Intelligent seismic data interpolation via convolutional neural network. *Progress in Geophysics*, 35(2):721–727, 2020. doi: https://doi.org/10.6038/pg2020DD0013.

Zhu, W. and Beroza, G. C. PhaseNet: a deep-neural-network-based seismic arrival-time picking method. *Geophysical Journal International*, 216(1):261–273, 10 2018. doi: 10.1093/gji/ggy423.

doi:10.26443/seismica.v3i1.1164

SEISMICA

Effects on a Deep-Learning, Seismic Arrival-Time Picker of Domain-Knowledge Based Preprocessing of Input Seismograms

Anthony Lomax *[1], **Matteo Bagagli** [2,3], **Sonja Gaviano** [4,5,7], **Spina Cianetti** [4], **Dario Jozinović** [6], **Alberto Michelini** [3], **Christopher Zerafa** [4], **Carlo Giunchi** [4]

[1]ALomax Scientific, Mouans-Sartoux, France, [2]Dipartimento di Scienze della Terra, Università di Pisa, Via Santa Maria, 53, 56126 Pisa, Italy, [3]Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Nazionale Terremoti, Via di Vigna Murata, 605, Roma, Italy, [4]Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Pisa, via Cesare Battisti, 53, Pisa, Italy, [5]Dipartimento di Scienze della Terra, Università degli Studi di Firenze, Via La Pira 4, Florence, Italy, [6]Swiss Seismological Service (SED), ETH Zurich, Zurich, Switzerland, [7]Now at [2]

Author contributions: Investigation: M. Bagagli, A. Lomax, S. Gaviano. Methodology: All authors. Software and Data Curation: M. Bagagli. Visualization: M. Bagagli, A. Lomax, S. Gaviano. Conceptualization and Supervision: A. Lomax, A. Michelini, C. Giunchi. Funding Acquisition and Resources: C. Giunchi, S. Cianetti, A. Michelini. Writing – original draft: A. Lomax. Writing – additions, review and editing: all authors..

Abstract Automated seismic arrival picking on large and real-time seismological waveform datasets is fundamental for monitoring and research. Recent, high-performance arrival pickers apply deep-neural-networks to nearly raw seismogram inputs. However, there is a long history of rule-based, automated arrival detection and picking methods that efficiently exploit variations in amplitude, frequency and polarization of seismograms. Here we use this seismological domain-knowledge to transform raw seismograms as input to a deep-learning picker. We preprocess 3-component seismograms into 3-component characteristic functions of a multi-band picker, plus modulus and inclination. We use these five time-series as input instead of raw seismograms to extend the deep-neural-network picker PhaseNet. We compare the original, data-driven PhaseNet and our domain-knowledge PhaseNet (DKPN) after identical training on datasets of different sizes and application to in- and cross-domain test datasets. We find DKPN and PhaseNet show near identical picking performance for in-domain picking, while DKPN outperforms PhaseNet for some cases of cross-domain picking, particularly with smaller training datasets; additionally, DKPN trains faster than PhaseNet. These results show that while the neural-network architecture underlying PhaseNet is remarkably robust with respect to transformations of the input data (e.g. DKPN preprocessing), use of domain-knowledge input can improve picker performance.

Production Editor: Gareth Funning
Handling Editor: Yen Joe Tan
Copy & Layout Editor: Théa Ragon

Signed reviewer(s): Mostafa Mousavi

Received: December 6, 2023
Accepted: June 2, 2024
Published: June 20, 2024

Riassunto Individuare l'arrivo delle fasi sismiche è fondamentale per il monitoraggio e la ricerca dei terremoti. Attualmente, la maggior parte dei programmi di riconoscimento (pickers) utilizza le deep neural network (DNN) su sismogrammi grezzi. Esistono però decadi di ricerche sul rilevamento automatico degli arrivi sismici basate su variazioni in ampiezza, frequenza e polarizzazione dei sismogrammi (domain-knowledge). Sfruttiamo queste conoscenze per pre-processare i sismogrammi grezzi in cinque serie temporali, ottenendo le tre funzioni caratteristiche di un picker multibanda, il modulo e l'inclinazione. Utilizzando questo nuovo input, realizziamo un'estensione di PhaseNet (PN) basata sulla domain-knowledge (DKPN) e confrontiamo i due modelli (PN e DKPN), addestrandoli su stessi dataset di diverse dimensioni. Eseguiamo due test: in-domain (su dati estratti dallo stesso dataset di addestramento) e cross-domain (su dataset diversi). DKPN e PhaseNet mostrano prestazioni quasi identiche per il riconoscimento delle fasi in-domain, mentre DKPN supera PhaseNet per alcuni casi cross-domain, in particolare per dataset di addestramento più piccoli. L'allenamento di DKPN è più veloce di quello di PhaseNet. Questi risultati mostrano che, sebbene l'architettura di rete neurale alla base di PhaseNet sia notevolmente robusta, l'uso di input basati sulla domain-knowledge può migliorare le prestazioni del picker.

Non-technical summary Automatic procedures for detecting seismic energy onsets on seismograms are critical for earthquake and environmental monitoring, earthquake and tsunami early-warning, and for fundamental research in seismology and earthquake hazard. Recent, high-performance onset detectors mainly use sophisticated, machine-learning algorithms which are "trained" on large sets of, unprocessed, seismograms. However, there is a long history of rule-based, automated onset detection algorithms in earthquake seismology that efficiently exploit various characteristics of seismogram waveforms. Here we use classical, seismological energy onset detection algorithms to transform seismogram waveforms before input to an established machine-learning onset-detector. We compare this extended detector with the original detector using identical training seismograms and application to diverse test seismograms. We find that the extended detector shows improved performance when applied to seismograms with different characteristics from those used for training, and can allow use of smaller datasets during training. The results show that the established machine-learning detector performs well independent of transformations of the input data, but that such transformations can improve performance and efficiency in some cases.

1 Introduction

Automated seismic arrival pickers are algorithms for detection, onset timing, phase type identification, and other characterization of seismic energy arrivals on seismogram waveforms. These pickers are fundamental for earthquake and environmental monitoring, earthquake and tsunami early-warning, arrival-time tomography, subsurface characterization, and for basic research of earthquakes and their hazard. Highly efficient, accurate and consistent automated picking is necessary for processing large to massive datasets with many data channels, high sampling rates or long recording periods, for real-time monitoring and warning, and for analyzing highly productive aftershock sequences and swarms.

For some years, automated seismic arrival picking algorithms have been developed using machine learning (Enescu, 1996; Dai and MacBeth, 1995; Wang and Teng, 1995; Mousset et al., 1996; Gentili and Michelini, 2006; Beyreuther et al., 2012; Kong et al., 2018). Recent, high-performance machine-learning pickers are based mainly on deep-neural-networks (LeCun et al., 2015) and are data-driven—trained and applied to nearly raw, seismogram waveforms as input features (Liao et al., 2021; Mousavi et al., 2019, 2020; Mousavi and Beroza, 2022; Münchmeyer et al., 2022; Ross et al., 2018a; Soto and Schurr, 2021; Woollam et al., 2019; Yu and Wang, 2022; Zhu and Beroza, 2018). There is, however, a long history of automated, seismic arrival detection, onset-timing and phase identification algorithms which efficiently exploit variations in amplitude, frequency and polarization of seismogram waveforms (Stevenson, 1976; Allen, 1978, 1982; Bai, 2000; Bagagli, 2022; McEvilly and Majer, 1982; Baer and Kradolfer, 1987; Withers et al., 1998; Lomax et al., 2012). These "classical" pickers are typically composed of rule-based algorithms defined by experts, using seismological domain knowledge to perform processing of seismogram waveforms, and subsequent analyses for detection, onset-timing and phase identification. Various parameters of the pickers are adjusted through trial-and-error or formal optimization (e.g., machine-learning Vassallo et al., 2012) to produce results best matching manually determined or other reference picks. Detections from classical pickers can also be fed into deep-neural-networks to further refine the arrival timing and characterize the picks (Yeck et al., 2020).

Incorporating expert, domain knowledge in the feature engineering and training of deep-neural-networks has been proposed and shown to improve performance over pure data-driven learning (Marcus, 2018; Borghesi et al., 2020; Jozinović et al., 2021; Kong et al., 2018; Mousavi and Beroza, 2022; Muralidhar et al., 2018), especially when there is limited or poor quality training data, or with difficult learning tasks. A basic question then arises: should the expert, seismological domain-knowledge components of classical pickers be discarded when developing high-performance, deep-neural-network pickers? This may be the case if the deep-neural-networks have a sufficiently large and complex architecture to learn to accurately map key characteristics of seismogram waveforms and phase onset energy into detections and picks, especially for cross-domain application. Such learning requires large, comprehensive and high-quality training datasets, and adequate computing resources to train the network. Otherwise, the use of domain-knowledge for preprocessing input seismogram waveforms may improve the performance, or even make viable, the training and application of deep-neural-network pickers, particularly for small or poor quality training datasets, when computing resources are limited, or for urgent analysis.

Most classical, automated pickers first convert raw seismograms into characteristic function (CF) time-series which greatly amplify the main features of seismic phase arrivals, such as abrupt changes in amplitude, frequency, or polarization of the waveforms. Secondly, these pickers analyze the CFs to detect and determine onset-times, phase types and other characteristics of possible seismic energy arrivals while ignoring background variations in signal. Conversion of raw seismograms into picker CFs is commonly and most basically performed with mean removal and high-pass filtering, followed by sliding-window calculation of a short-term average (STA) and a long-term average (LTA) of the signal amplitude to form the CF based on the ratio STA/LTA (Allen, 1978, 1982; Baer and Kradolfer, 1987). Additional or alternative processing of seismograms for detection, time picking or phase identification include autoregressive analysis (Sleeman and van Eck, 1999), particle-motion and polarization analysis (Vidale, 1986; Bai, 2000; Plešinger et al., 1986; Anant and Dowla, 1997; Ross and Ben-Zion, 2014), vector modulus (Bai, 2000), and time-frequency domain, spectrogram (Lomax et al., 2012; Alvarez et al., 2013; Njirjak et al., 2022) or wavelet analyses (Anant and Dowla, 1997; Zhang et al., 2003; Mousavi et al., 2016). FilterPicker (Lomax et al., 2012) constructs a picker CF through application of an STA/LTA algorithm to a series of band-pass filtered seismograms, equivalent to a simplified spectrogram representation of the raw waveforms. The multi-band nature of FilterPicker enables picking of seismic onsets over a range of dominant frequencies in the presence of signal offset and high noise, improving correct detection of true seismic phase onsets even in complex waveforms.

Here we examine changes in the performance of a deep-learning picker when its raw seismogram input is modified using seismological domain-knowledge from classical pickers. In a manner similar to Gentili and Michelini (2006) for picking and Wang and Teng (1995) and Njirjak et al. (2022) for earthquake detection, we preprocess 3-component, broadband seismograms into a set of 5 input time-series: the 3-component characteristic functions of the multi-band FilterPicker, plus the instantaneous modulus and inclination of the waveforms from particle-motion analysis. This preprocessing increases the dimensionality of the input data. We use these 5 time-series as input features instead of 3-component, raw seismograms to extend the deep-neural-network picker PhaseNet (Zhu and Beroza, 2018) within the SeisBench platform (Woollam et al., 2022).

*Corresponding author: anthony@alomax.net

We compare the original PhaseNet and our extended, domain-knowledge PhaseNet (DKPN) using identical training, validation and test datasets and identical processing pipelines. We train PhaseNet and DKPN on waveforms from the INSTANCE dataset (Michelini et al., 2021). The training is run on 7 subsets with different sizes, leading to a set of trained model variants. We apply this set of trained models to in-domain test datasets from INSTANCE and to cross-domain test data from two different datasets available in SeisBench: ETHZ from the Swiss Seismological Service, and PNW from seismic networks in the US Pacific Northwest (Ni et al., 2023).

The use of domain-knowledge input in DKPN instead of near-raw waveforms for the PhaseNet deep-learning picker requires slightly more computing time for training and application, due to the additional waveform preprocessing, though data preprocessing, code optimization, multi-processing and use of GPUs effectively removes this time penalty. For P and S arrivals the explicit information targeting detection and picking of seismic energy arrivals introduced by the domain-knowledge processing enables DKPN to reach higher performance than PhaseNet for cross-domain application, especially with smaller training datasets. In contrast, DKPN and PhaseNet perform nearly identically for in-domain P and S picking. These results suggest that the underlying PhaseNet architecture can robustly learn arrival detection and pick characterization somewhat independently of the form of the input data, presumably as long as key information relevant to arrival detection and picking remains present in the input, as is the case with DKPN.

2 Data and Methods

We use the SeisBench machine learning toolbox (Woollam et al., 2022) to access seismogram waveform datasets and the PhaseNet deep-neural-network picker model, and as a general platform for data processing and augmentation operations.

2.1 PhaseNet, a deep-neural-network picker

PhaseNet (Zhu and Beroza, 2018) is a deep-neural-network algorithm for probabilistic detection, onset-timing and phase-type identification of seismic P and S arrivals.A trained deep-neural-network can be interpreted as a very high-dimensional approximation function composed of many, local mappings of input to output (Balestriero and Baraniuk, 2018). These mappings produce increasingly abstract layers which preserve only essential information in the data needed for a target regression or classification task (LeCun et al., 2015). Indeed, in this study we are investigating the effects of using essential information for picking as input and thus potentially reducing the amount of network training needed to identify and isolate essential information.

The input for PhaseNet are 3-component, broadband, seismogram waveforms of 3001 samples (30 sec at 100 Hz sampling) with minimal preprocessing (mean removal and normalization). These input data are processed through a modified U-net (Ronneberger et al., 2015; Zhu and Beroza, 2018, , their figure 5) with 4 stages of down-sampling and reduction in number of nodes based on 1-D convolution followed by 4 stages of near-symmetric up-sampling and expansion based on 1-D deconvolution (Zhu and Beroza, 2018). Direct, skip connections between corresponding down- and up-sampling layers help to improve training convergence. After the last stage of down-sampling, the input is reduced to 22 points x 12 channels, which, considering the size of the convolutional kernel , implies the network has a broad receptive field on the original seismograms (Zhu and Beroza, 2018) which is about 26 sec (Hien, 2018). At the end of up-sampling, PhaseNet outputs 3 channels of 3001 points: prediction probability distributions for P and S arrivals and for noise, time-aligned to the original 3001 input samples. Here we derive arrival picks from peaks in the P and S probabilities through rule-based post-processing to give pick arrival time (at peak maximum), and confidence (peak amplitude)

In Zhu and Beroza (2018) PhaseNet is trained on a dataset of detected earthquakes from Northern California composed of 623,054 3-component recordings, all of which have manually picked, P and S arrival times. Through various experiments, Zhu and Beroza (2018) conclude that PhaseNet achieves much higher picking accuracy and recall rate than a classical STA/LTA plus autoregressive analysis method (Akazawa, 2004) when applied to the waveforms of known earthquakes, with a particularly pronounced improvement in S picking performance.

2.2 Modified FilterPicker characteristic functions

FilterPicker (Lomax et al., 2012) applies an adaptation of the STA/LTA algorithm of Baer and Kradolfer (1987) to construct a picker CF from a succession of band-pass filtered seismograms. Seismogram waveforms, with little or no preprocessing, pass through a pipeline of: 1) differentiation, 2) band-pass filtering at a geometric progression of center periods ranging from the sampling interval to the longest period signal to be picked (e.g., for sampling interval 0.01sec, 7 bands at 0.01, 0.02, 0.04, 0.08, 0.16, 0.32, 0.64 sec center periods), 3) squaring of each band-pass series to form a positive, envelope function, 4) forming a CF for each band as the ratio of instantaneous deviation from long-term mean to long-term standard deviation of the band envelope, and 5) forming a definitive, summary CF from the maximum of the band CFs at each sample point. The succession of band-pass filtered seismograms are equivalent to a simplified time-frequency, spectrogram representation of the raw waveforms. The multi-band nature of FilterPicker provides a strong response in the summary CF for seismic onsets with a range of dominant periods, even in the presence of strong, narrow-band noise, of strong micro-seismic or other, longer period noise, and of offset signals. See Lomax et al. (2012) for details and examples.

We implement FilterPicker within the data augmentation step of SeisBench processing using the FBpicker, Python implementation of FilterPicker from

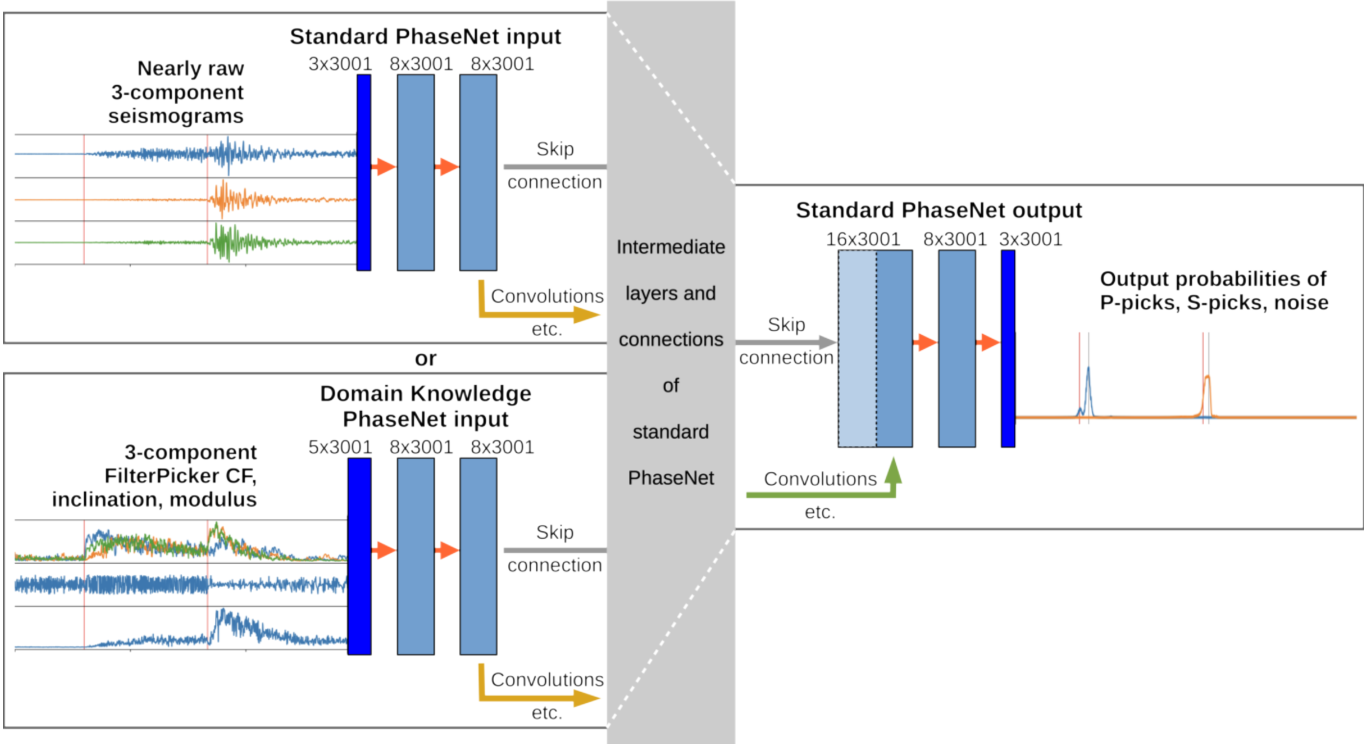


Figure 1 Standard PhaseNet and modified Domain Knowledge PhaseNet deep-learning architecture. The DKPN model replaces the nearly-raw, 3-component seismogram input of PhaseNet with the 3-component FilterPicker CF plus instantaneous modulus and inclination traces (bottom left panel). Otherwise, the layers, connections and output for both models are identical. For details of the full PhaseNet architecture, symbols and color codes, see Zhu and Beroza (2018, ; their figure 4).

the PhasePApy package (Chen and Holland, 2016). We modify the FBpicker algorithm to replace its sliding, fixed-window procedure for generating band CFs with the decay-constant, recursive filter procedure of the original FilterPicker. We further modify the resulting FilterPicker algorithm by taking the logarithm of the summary CF to compress high amplitudes in the CF at strong arrival onsets and thus avoid the need for a cut-off parameter (Lomax et al., 2012) to limit the highly-variable maximum CF values. Finally, we normalize the CFs with the maximum standard-deviation of the 3-component CFs. As we do not use the detection and pick characterization logic of FilterPicker, there are only two primary picker parameters: a filter window defining the frequency of the longest period band, and the long-term, time-averaging scale for recursive band-pass filtering. In this study we set FilterPicker parameters following the guidelines and defaults in Lomax et al. (2012), with some trial-and-error over a limited range of typical values for broadband, local and regional event picking.

2.3 Instantaneous modulus and inclination

In addition to the modified, 3-component CF waveforms from FilterPicker, we also calculate two waveforms consisting of quantities from particle-motion analysis, the instantaneous modulus and inclination, using the FilterPicker, band-pass filtered seismograms. In order to suppress response to background noise, both quantities are calculated independently at each sample point, using the 3-component Z, N, E data values on the band-pass waveform corresponding to the maximum band CF for the sample point. The modulus is the length of the 3-component data vector, $\sqrt{(Z^2 + N^2 + E^2)}$, normalized by dividing by the maximum standard-deviation over all sample points. The inclination or incidence angle is given by $\tan^{-1}[Z / \sqrt{(N^2 + E^2)}] / \pi$, where dividing by π normalizes to a range of [-1,1] so -1, 0 and 1 correspond to down, horizontal and up inclination, respectively.

We expect these two additional waveform inputs will help the picker to recognize and discriminate between P and S phases in the stream, especially given that the replacement of raw waveforms with CF's entails a loss of information. Modulus re-introduces absolute amplitude information which might aid in discrimination of, for example, a typically higher amplitude S arrival from a lower amplitude P arrival. Inclination re-introduces polarization information which can indicate the type of arrival since P waves usually exhibit dominantly vertical particle motion and S waves often show stronger horizontal motion.

2.4 Domain Knowledge PhaseNet

DKPN is our extension of the PhaseNet picker model to use classical, seismological domain-knowledge as input features. We replace the 3-component, raw seismogram input to the PhaseNet model with the 3-component CFs of the modified FilterPicker and the instantaneous modulus and inclination of the waveforms (Figures 1 and 2-5). This entails that in the first convolution plus rectified linear unit step in the first layer, instead of transforming a 3x3001 input dimension to 8x3001 dimension

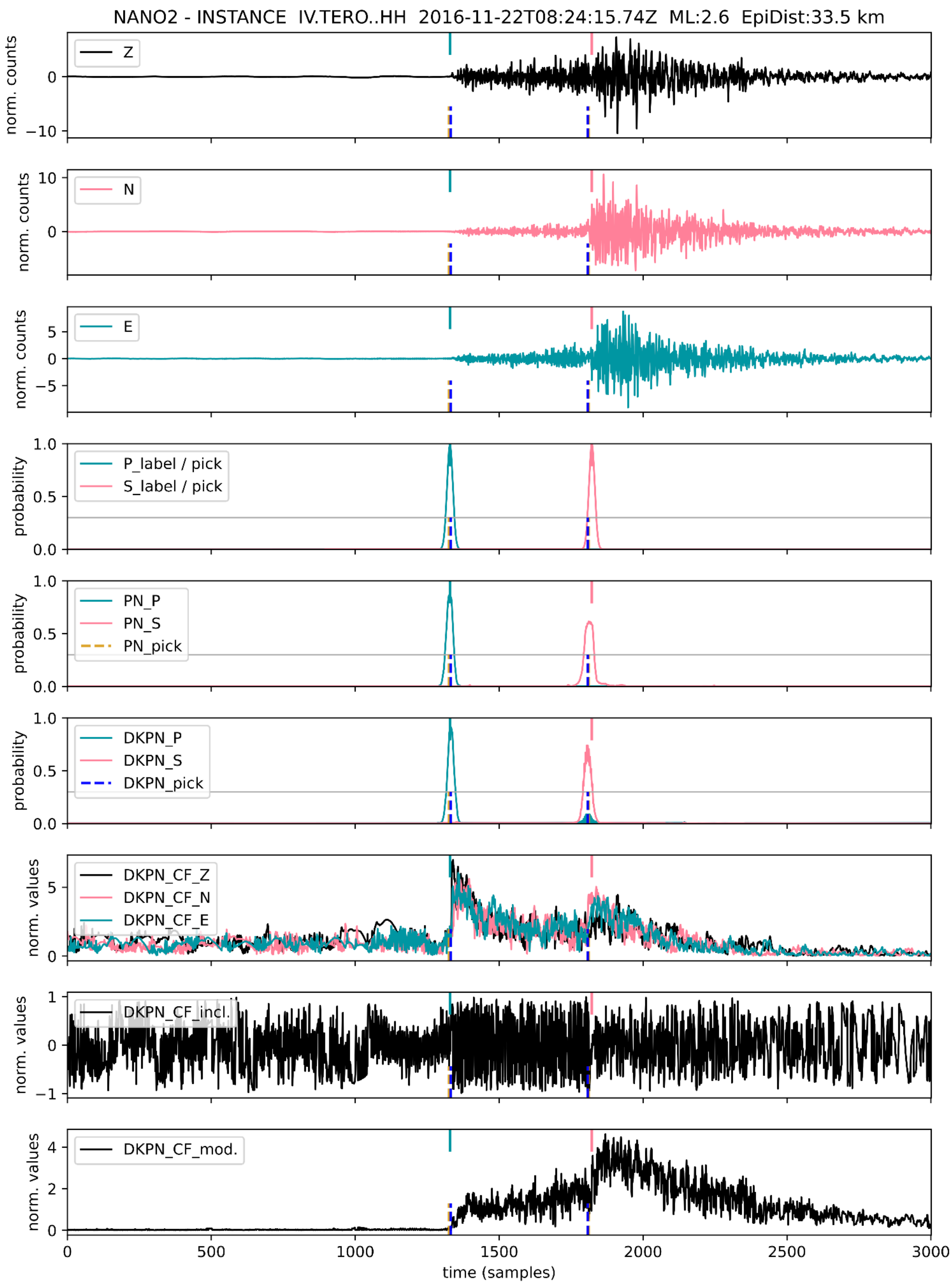


Figure 2 Example processing and picking results for models trained with the INSTANCE NANO2 dataset and applied to INSTANCE, ETHZ, and PNW trace samples. In each plot, the title indicates INSTANCE training dataset size, test dataset, network-station-location-channel code, event date, magnitude and distance from station. Subplots show: (rows 1-3) de-meaned and normalized Z, N, and E component, observed seismograms; (row 4) P (blue-green) and S (light red) label data; (rows 5-6) PhaseNet (PN) and DKPN P (blue-green) and S (light-red) probabilistic predictions, gray horizontal lines indicate threshold 0.3; (row 7) normalized, DKPN FilterPicker Z, N, E CF's; (rows 8-9) normalized, DKPN instantaneous inclination and modulus. Solid bars at top of each subplots show P (blue-green) and S (light red) label picks. Dashed bars at the bottom of each subplot show PhaseNet (yellow) and DKPN (blue) P and S predicted picks for threshold 0.3. Horizontal axis shows sample count. Seismogram from the INSTANCE dataset (in-domain testing) for which DKPN and PhaseNet both pick P and S strongly near the label times. Note the sharp, strong P onset and emergent S onset in the DKPN CFs, the change in character between P and S of the DKPN inclination, and the clear P and S onsets in the DKPN modulus; these features show the introduced domain knowledge which drives the DKPN picks.

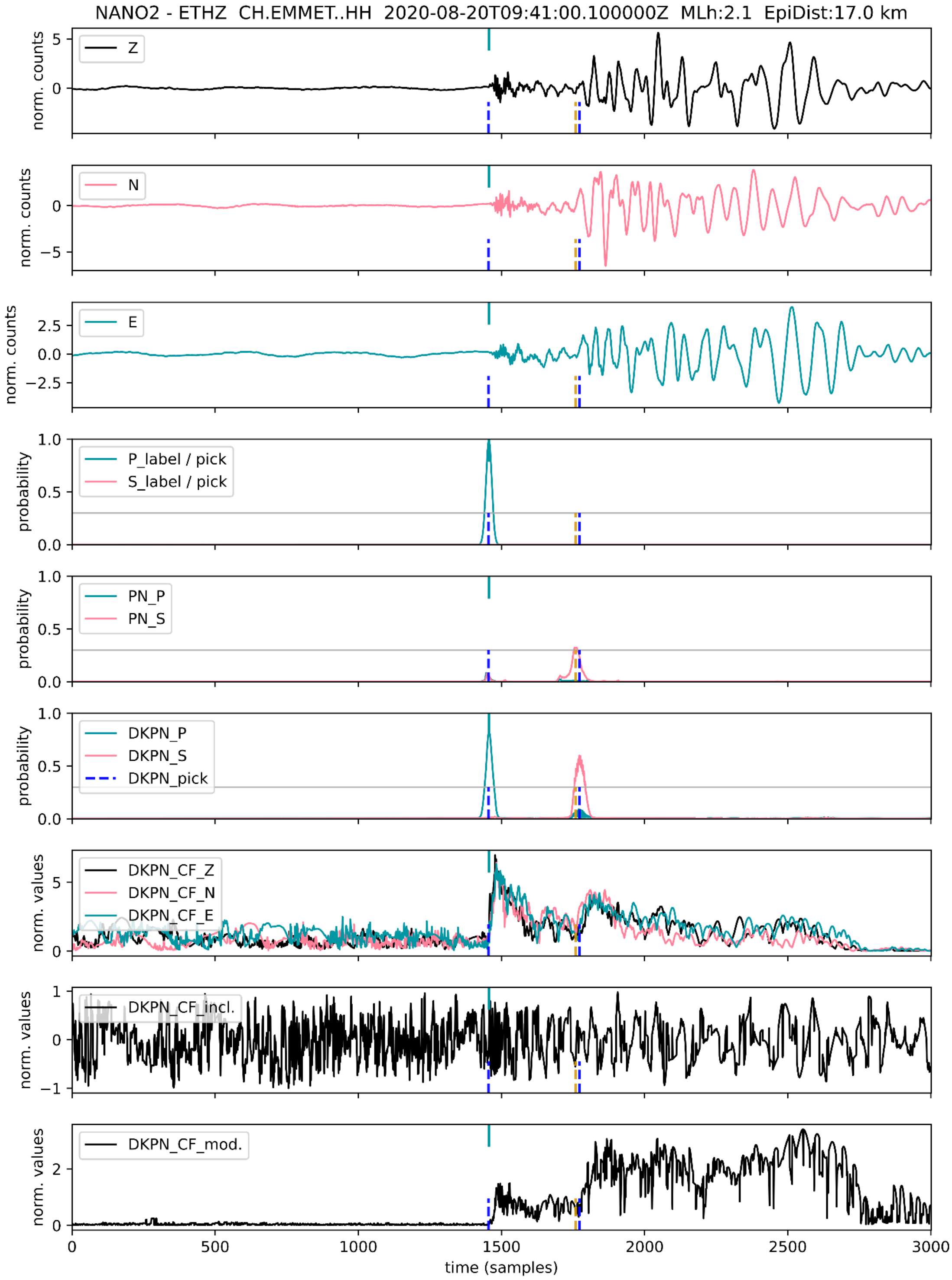


Figure 3 Same as Figure 2, for a seismogram from the ETHZ dataset (cross-domain testing) for which a clear, high-frequency, labeled P arrival is matched by DKPN but not PhaseNet. DKPN strongly, and PhaseNet more weakly both identify an unlabeled S arrival, which is likely correct given the waveforms and epicentral distance. The relatively poor performance of PhaseNet for this seismogram may be due to the waveform arrivals (e.g. short, high-frequency P signal and long duration, long-period S signal) differing significantly from arrival waveforms in the INSTANCE dataset used for training.

as in standard PhaseNet (Zhu and Beroza, 2018, ; their figure 4), DKPN transforms a 5x3001 input dimension (channels x length) to 8x3001 dimension (Figure 1). We otherwise make no change to the PhaseNet model as implemented in SeisBench (SeisBench v0.3 or later).

2.5 Seismogram waveform datasets

In this study we use three benchmark, seismogram waveform datasets provided in SeisBench: INSTANCE (Michelini et al., 2021) composed of ~1.2 million, 3-component waveforms for ∼50,000 earthquakes from M 0 to M 6.5 in and around the Italy region (epicentral distances ~0-6°); Eidgenössische Technische

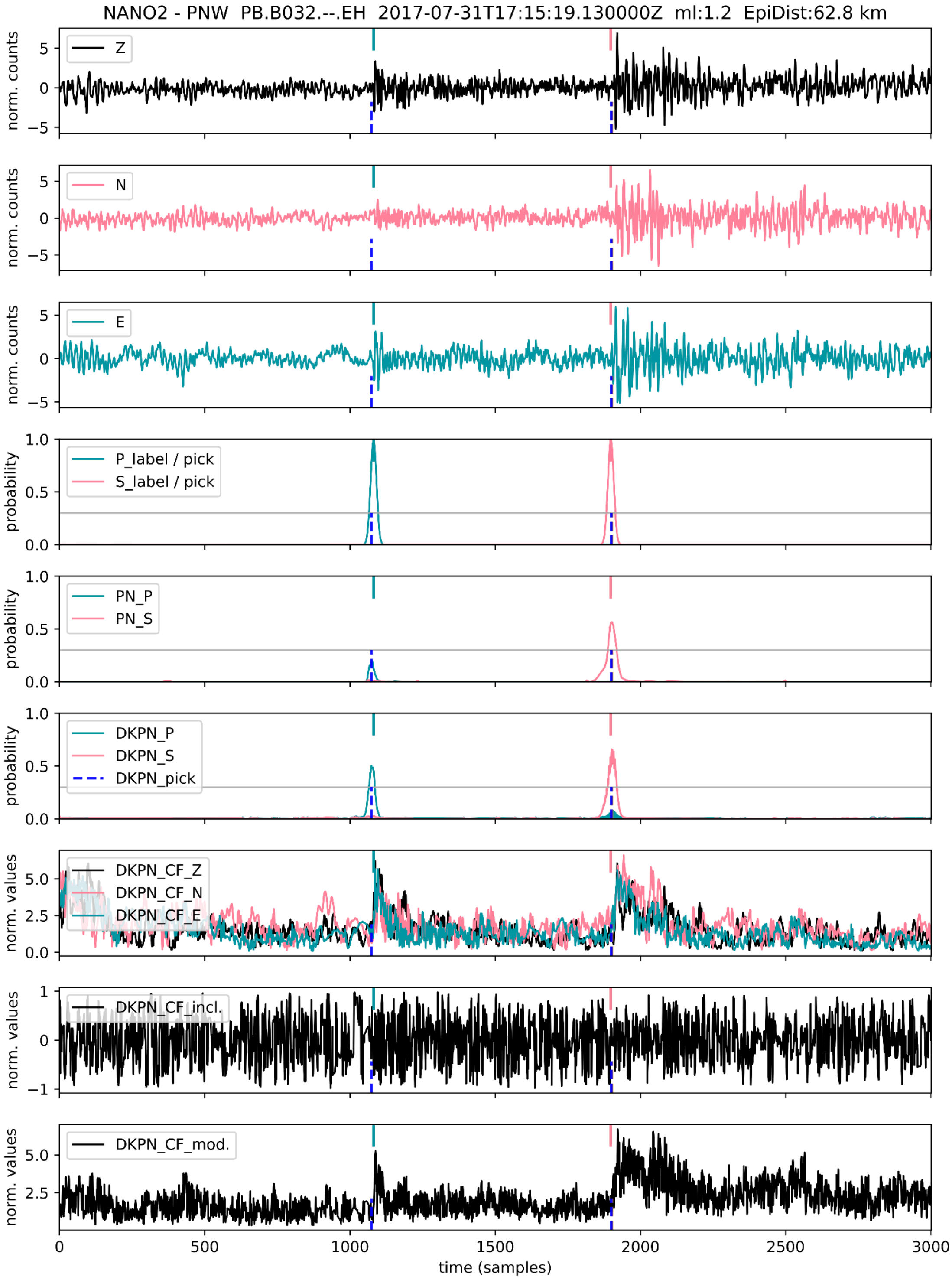


Figure 4 Same as Figure 2, for a low S/N seismogram from the PNW dataset (cross-domain testing) for which DKPN correctly picks a P arrival with a sharp prediction probability peak, while PhaseNet fails to pick but shows a weak prediction probability peak.

Hochschule Zürich (ETHZ; Woollam et al., 2022) composed of 36,743, 3-component waveforms for 2231 events from M 1.5 to ~M 5 in and around the Switzerland region (epicentral distances ~0-4°); and the Pacific Northwest AI-ready Seismic Dataset (PNW, Ni et al., 2023) composed of ~200,000, 3-component waveforms for ~65,000 events from M 0 to ~M 6.4 in the US Pacific Northwest region with local and regional epicentral distances. A signal-to-noise (S/N) ratio included in the INSTANCE metadata, reported in dB, is calculated (Michelini et al., 2021) from amplitudes in the 5 sec following the S arrival (signal) and 5 sec before the P arrival (noise), though here we do not filter the INSTANCE training datasets on S/N ratio. INSTANCE includes pure noise waveforms, which, following Zhu and Beroza (2018) we do not use for training. For detailed information on filtering criteria for both training and testing stages, the reader is referred to the Supplementary

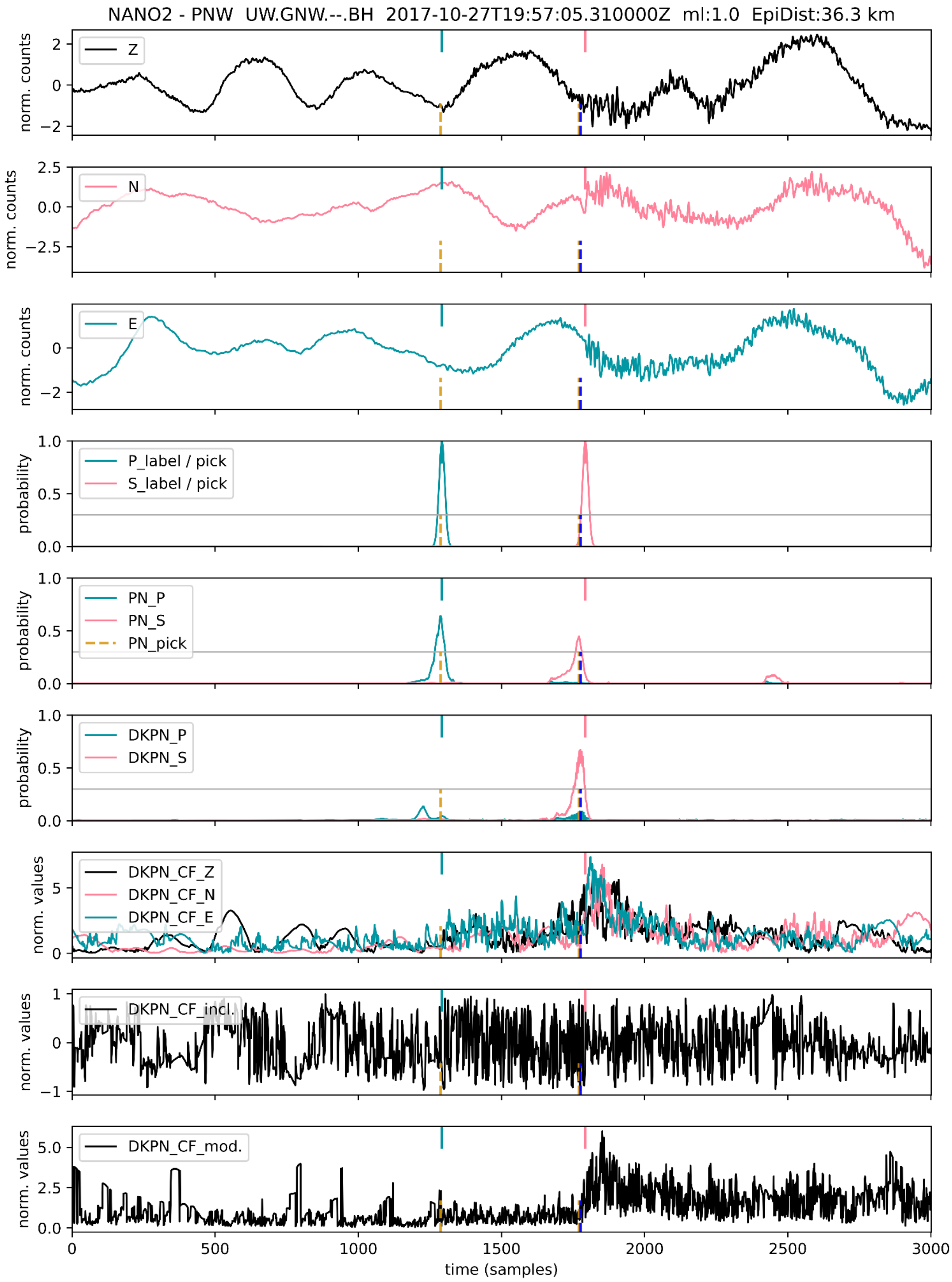


Figure 5 Same as Figure 2, for a low S/N seismogram from the PNW dataset (cross-domain testing) for which PhaseNet correctly picks a weak P arrival with a sharp prediction probability peak, perhaps by responding to the low amplitude, P coda signal. DKPN fails to pick P, with a very weak prediction probability peak; this failure is likely due to the emergent CF Z component and lack of amplitude increase on the modulus at the P time, both related to the P arrival on the Z seismogram having low amplitude and frequency content similar to the preceding noise.

Text S1 and Table S1.

Visual examination of INSTANCE dataset waveforms and picks shows some traces with: missing S picks; labeled picks on very high or pure noise low-gain data; clearly early or late picks; and unreasonably large pick uncertainty estimates. To mitigate these problems, we filter the dataset metadata to include only high-gain HH channels, events at epicentral distance ≤ 100 km, and, following Zhu and Beroza (2018), traces which have both P and S picks, resulting in ~300,000 3-component trace sets available for training, validation and testing. We also create probabilistic pick labels for training using a fixed sigma of 0.1 sec instead of using the labeled pick uncertainties. See Supplementary File S1 for examples of filtered INSTANCE waveforms.

In order to evaluate picker performance when different amounts of training data are available, we establish INSTANCE training datasets with different sizes ranging from insufficient for training convergence and stability (NANO3, ~900 samples), to minimal for convergence (NANO2, ~1.6k samples), through intermediate sizes (NANO1, ~3k samples; NANO, ~6k samples; MICRO, ~12k samples; TINY, ~24k samples; SMALL, ~61k samples; MEDIUM, ~163k samples) to much larger than needed for apparent convergence for both pickers (LARGE, ~245k samples). In the following we focus on results for the NANO2, MICRO and MEDIUM training datasets as representative of the main evolutions and features of picker performance across all training dataset sizes.

Missed S picks are also a problem with the ETHZ dataset, but apparently much less so with the PNW dataset. As we do not train on these datasets we do not remove these potential problem traces (examples of used ETHZ and PNW waveforms are shown in Supplementary Files S2 and S3, respectively). However, due to missed S picks, our testing statistics and metrics for S for the ETHZ dataset are likely degraded. Similar and other quality problems are likely present in most seismological waveform datasets for machine-learning, as Münchmeyer et al. (2022) discuss for the datasets provided in SeisBench. Such errors in the data labels will adversely affect the rate and quality of model training and bias the validation and test statistics and metrics, but similarly for the two picker algorithms, thus highlighting their performance in practice.

FilterPicker, like most STA/LTA pickers, requires a minimum length of background data before any phase arrivals for statistical stabilization; this length is controlled by the long-term window parameter. For machine learning training in general, an ample length of background before arrivals is also needed for random window-shift, data augmentation to enhance generalization in the trained model, and, most importantly, to avoid that first arrivals are near the same window position in all or most training samples. Lack of sufficient background data before arrivals can impair classical methods like STA/LTA in comparisons with machine-learning pickers. The processing workflow described below addresses this required minimum length of background data. We note also that lack of sufficient background data before arrivals can preclude pertinent training and evaluation of machine-learning pickers for real-time and early-warning application, where in practice an almost unbounded amount of data before an arrival is available, and very little data may be available after an arrival onset before reaching the last received data sample.

2.6 Dataset and model configuration, processing, training and comparison

To compare the performances of PhaseNet and DKPN on different test sets we configure and preprocess datasets and models, train the models and compare PhaseNet and DKPN P and S arrival predictions, statistics and metrics (see Data and code availability). Care must be taken at the data generation stage to ensure that the waveform time-series has sufficient length before the first label pick for FilterPicker stabilization (FPS); this length should be greater than the number of sample points (N_{FPS}) corresponding to the FilterPicker long-term, time-averaging scale. Figures 2-5 show preprocessing, label and pick prediction waveforms for trace samples from the INSTANCE, ETHZ and dataset.

The configuration and processing workflow includes:

1. Load the requested dataset; set the sampling rate to 100 Hz.
2. Optionally select or mask dataset trace sets on presence of P and S picks, channel code, epicentral distance, or other available meta-data. (Supplementary Text T1)
3. For training, randomly split the data into training, validation and test sets according to a target training set size (e.g. for our INSTANCE MEDIUM training dataset: 50% training, 5% validation and 45% remainder for drawing test samples).
4. Define data generators with identical preprocessing and augmentation steps for training, validation and testing, the principal steps are:
    (a) Get a randomly positioned data window in the input, 3-component seismograms starting at least N_{FPS} points before the first pick label, and with a length of N_{FPS} plus the 3001 points required for input to the deep-learning models.
    (b) Normalize to the maximum standard-deviation across the 3-component data.
    (c) For DKPN, apply the processing described in the section "Modified FilterPicker characteristic functions" to generate the 3-component, FilterPicker CF time-series, and apply the processing described in "Instantaneous modulus and inclination" to generate the modulus and inclination time-series. In this study the FilterPicker filter window is set so the longest period signal analyzed is 2 sec, and the long-term, recursive-filter time-averaging scale to 4 sec, giving $N_{\mathrm{FPS}} = 401$ point.
    (d) For DKPN, cut the first N_{FPS} (after FilterPicker stabilization) for all time-series to form a data window of length 3001 points as required for input to the deep-learning models.
    (e) Create probabilistic, P and S label traces with the same 3001 point window from picks specified in the trace metadata using the probabilistic labeler function in SeisBench. Each P or S pick is summed into the corresponding P or S trace as Gaussian function of amplitude 1.0 and with a fixed variance of 5 points (0.05 sec for 100 Hz data).

The training workflow includes:

1. Model and dataset setup following steps 1-4 of the configuration and processing workflow.

2. Run the training on the training dataset for specified optimizer and loss functions, learning rate and number of epochs. In this study the Adam optimizer (Kingma & Ba, 2017), and a cross-entropy loss function are always used. Train models using an early-stopping approach governed by a *"patience"* parameter indicating the number of epochs to tolerate without improvement, and a fixed "*improvement*" threshold. If the mean validation loss over the last two *patience* length epochs does not exceed the *improvement* threshold compared to the mean development loss over the preceding *patience* epochs, the training process is halted. A second condition for halting the training process is if the validation loss over the last *patience* epochs consistently surpasses the training loss, indicating a potential overfitting tendency. In either halting case, the model weights obtained *patience* epochs before the current epoch are utilized. We use this approach to respect the different, natural learning-time behavior of the 2 algorithms, and to smoothly converge to the best possible minima. For details of the training parameters and loss curves comparisons see Supplementary Text S1, Table S1, and File S4).

The comparison workflow includes:

1. Model and dataset setup following steps 1-4 of the configuration workflow.

2. Load a trained model and apply it to traces drawn from the test dataset to obtain prediction probability distributions for P and S arrivals.

3. Post-process the probabilistic, P and S prediction traces with:

   (a) 3-point smoothing to suppress rapid oscillation,

   (b) pick detection at peaks of amplitude greater than specified thresholds and separated by more than 0.5 sec,

   (c) retain the pick time, T_p and amplitude, A_p.

4. Accumulation of the P and S, Gaussian labels and prediction picks for multiple traces from the test dataset for calculation of evaluation statistics and metrics as described in the following.

In this study we repeat the training workflow using 7 different, randomly selected training and validation subsets of traces for each experiment. We therefore obtain 7 different models and 7 sets of test results for each individual training dataset size. We merge these test results to reduce the dependence of testing performance statistics with respect to training dataset selection.

2.7 Evaluation statistics and metrics

As a basis for comparison of PhaseNet and DKPN performance on test datasets, and following (Zhu and Beroza, 2018), we count, relative to the labeled P or S arrivals in the test dataset, the number for P or S of correct Gaussian predicted arrivals (true positives; TP), incorrect predicted arrivals (false positives; FP), and no prediction of a labelled arrival (false negatives; FN). There may be multiple FP picks for P or for S on each data window. Here, a smoothed, Gaussian predicted P or S arrival is counted as correct when its peak amplitude, A_p, is greater than a specified threshold and the difference, ΔT_p, between its peak time and the time of a label arrival of the same phase is less than 0.1 sec for P and 0.2 sec for S, which has typically noisier onsets than P arrivals. Zhu and Beroza (2018) use $\Delta T_p \leq 0.1$ for P and S, and use a threshold $A_p \geq 0.5$. In this work, we examine a range of thresholds $0.1 \leq A_p \leq 0.9$ to find optimal metrics such as F1 score, which vary with training dataset size, test dataset and for PhaseNet versus DKPN. While use of a low amplitude threshold leads to a higher rate of picking, we find that low amplitude predictions generally correspond to correct arrival picks. Additionally, filtering of a limited number of false picks can be done in phase association and hypocenter location processing stages (Kim et al., 2023). For advanced hypocenter location and other algorithms (e.g., Satriano et al., 2008; Lomax et al., 2014) the peak amplitude A_p can also be used for weighting or selection of picks and some measure of the width of the peak (e.g. at half its height) as a pick uncertainty.

From the TP, FP, FN statistics, we form the following metrics for P and for S:

Precision, the proportion of positive arrival predictions that are correct,

$$P = \frac{TP}{TP + FP} \tag{1}$$

Recall, the proportion of labeled positives that are correctly predicted,

$$R = \frac{TP}{TP + FN} \tag{2}$$

and the *F1 score*, which balances the often opposing, Precision and Recall metrics through their harmonic mean,

$$F1 = 2x\frac{PxR}{P + R} \tag{3}$$

3 Results

We present a set of tests to show and compare the performance of PhaseNet and DKPN for different training dataset sizes applied to in-domain (INSTANCE) and cross-domain (ETHZ and PNW) test datasets.

3.1 Test 1 – In-domain

We first compare the in-domain performance of PhaseNet and DKPN across different training dataset

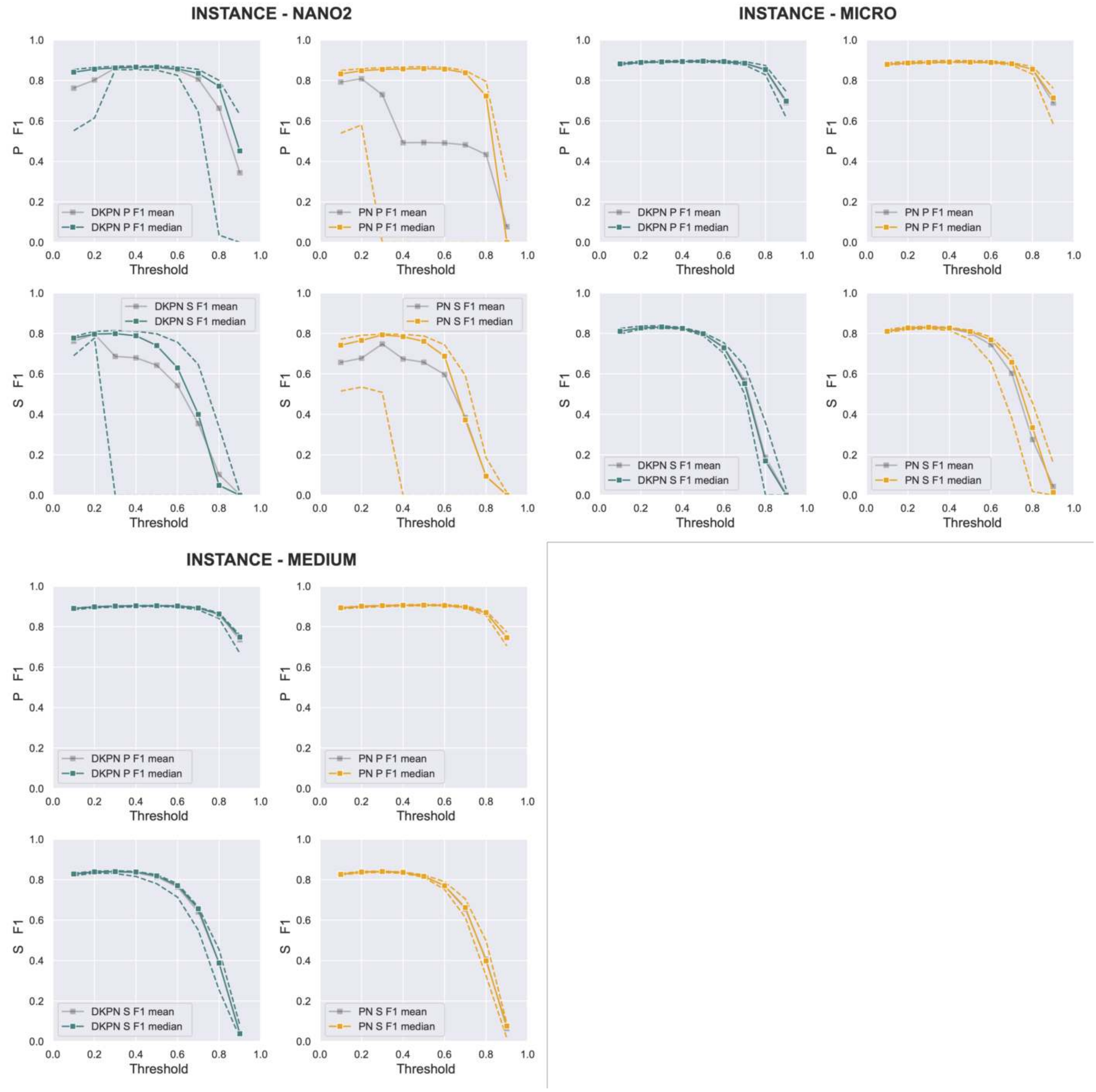


Figure 6 Comparison of PhaseNet (PN) and DKPN with testing on the INSTANCE dataset (in-domain). F1 score metrics across a range of pick amplitude thresholds for P arrivals and S arrivals for models trained with different size INSTANCE datasets. Mean, median and upper/lower limits of the mean (dashed curves) of F1 for 7 runs each with 5,000 evaluation samples drawn from the test datasets, these samples vary for each training dataset size.

sizes and amplitude thresholds to define correct predicted picks. We train and test with data samples from the INSTANCE dataset. Sample INSTANCE waveforms are presented in Supplementary File S1. The F1 metrics for this test are shown in Figure 6.

For the larger training datasets (e.g. MICRO and MEDIUM) DKPN and PhaseNet show almost identical performance, with F1 scores of about 0.9 for P up to a threshold of about 0.7 and F1 about 0.8 for S up to thresholds of about 0.4 - 0.5. The reduced performance of both pickers for S is almost certainly due to difficulties for both models in detecting and picking the S onset, which is often embedded in the P coda and emergent.

For the smaller NANO2 training dataset, the DKPN and PhaseNet, median P and S F1 scores are slightly reduced relative to those with the larger training datasets. The mean and lower limit of F1 scores, however, show a significant degradation, likely indicating instability in training with small datasets and chance of convergence to an inadequately trained model, even for application to an in-domain dataset.

Histograms of P and S pick residuals (predicted time - label time) for PhaseNet and DKPN for a selection of training data set sizes are shown in Figure 7. For P the total number of predicted picks within twice ΔT_p is generally similar for PhaseNet and DKPN and independent

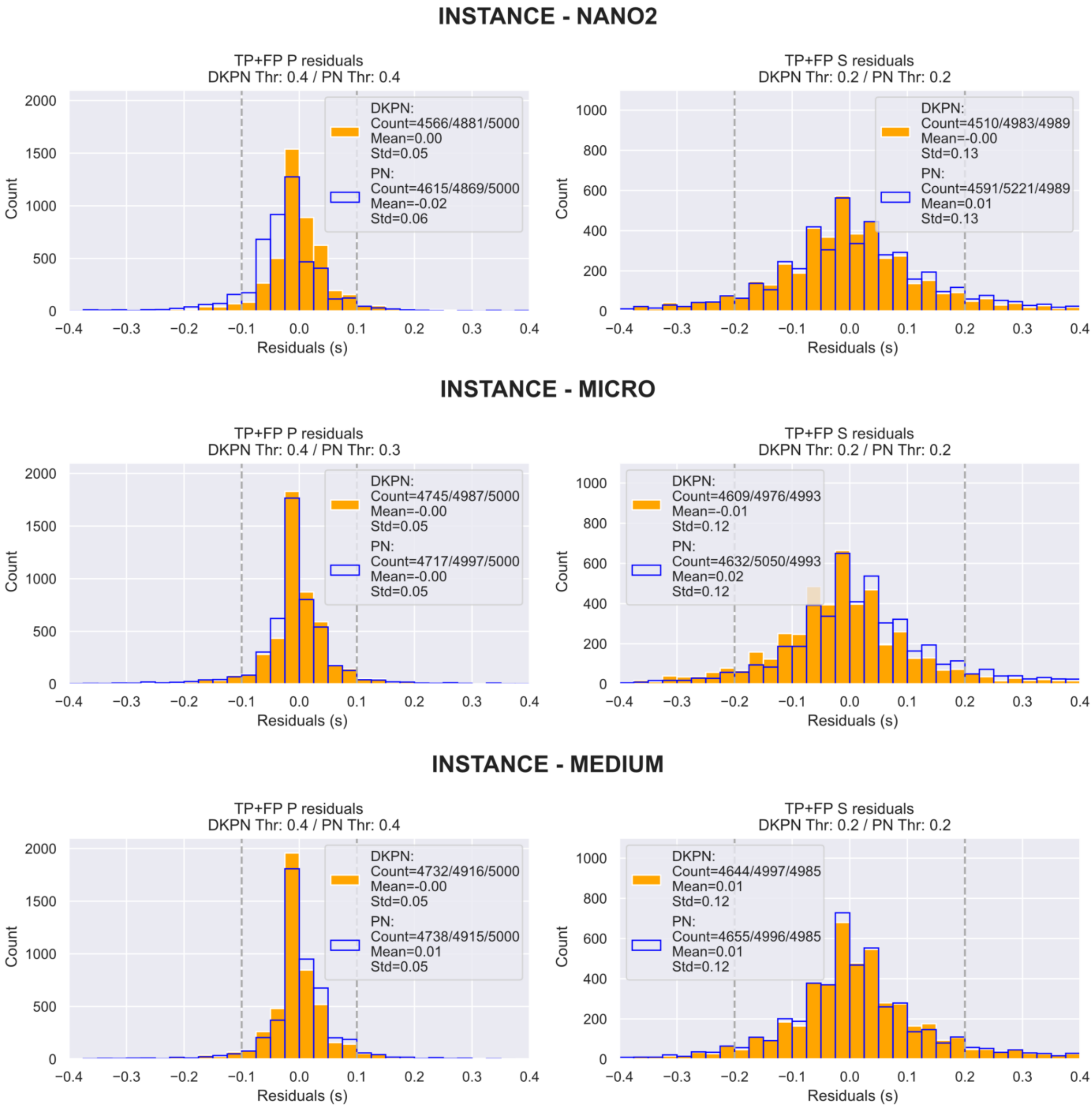


Figure 7 Histograms of P and S pick residuals for selected INSTANCE training dataset sizes tested on the INSTANCE dataset (in-domain). Results shown for the pick amplitude threshold (Thr) giving the highest F1 score for each case of dataset size and method (DKPN or PhaseNet). Vertical, dashed gray lines show the maximum difference ΔTp between a pick time and the corresponding label time (0.1 sec for P and 0.2 sec for S) to declare a correctly predicted arrival for evaluation statistics (TP). Pick counts show: Number of residuals (number of predicted picks) used in the mean and standard-deviation statistics / Total number of residuals available / Total number of label picks available for the test case. To remove outlier data, the mean and standard-deviation statistics use trimmed residuals, within twice ΔTp: ± 0.2 sec for P and ± 0.4 sec for S.

of training dataset size, and there is little variation in the mean of the residuals, which is always near zero, or for standard deviation with training dataset size. However, the DKPN P residuals are slightly more concentrated and peaked around zero than are the PhaseNet residuals for NANO2. For S the distribution of residuals and statistics for the two pickers are very similar and show little variation with training dataset size, aside from slightly more concentration of residuals around zero for the larger training dataset sizes (e.g. MEDIUM).

Figures 2-5 show test results for PhaseNet and DKPN models trained with the INSTANCE NANO2 dataset and applied to INSTANCE, ETHZ and PNW testing trace samples. These examples illustrate how the FilterPicker CFs, inclination and modulus relate to and help determine the probabilistic P and S predictions and picks for DKPN, and how the amplitude and complexity of probabilistic predictions for PhaseNet and DKPN relate to the trace noise and to the complexity and impulsiveness of arrival onsets. Note in particular how the DKPN CF's for

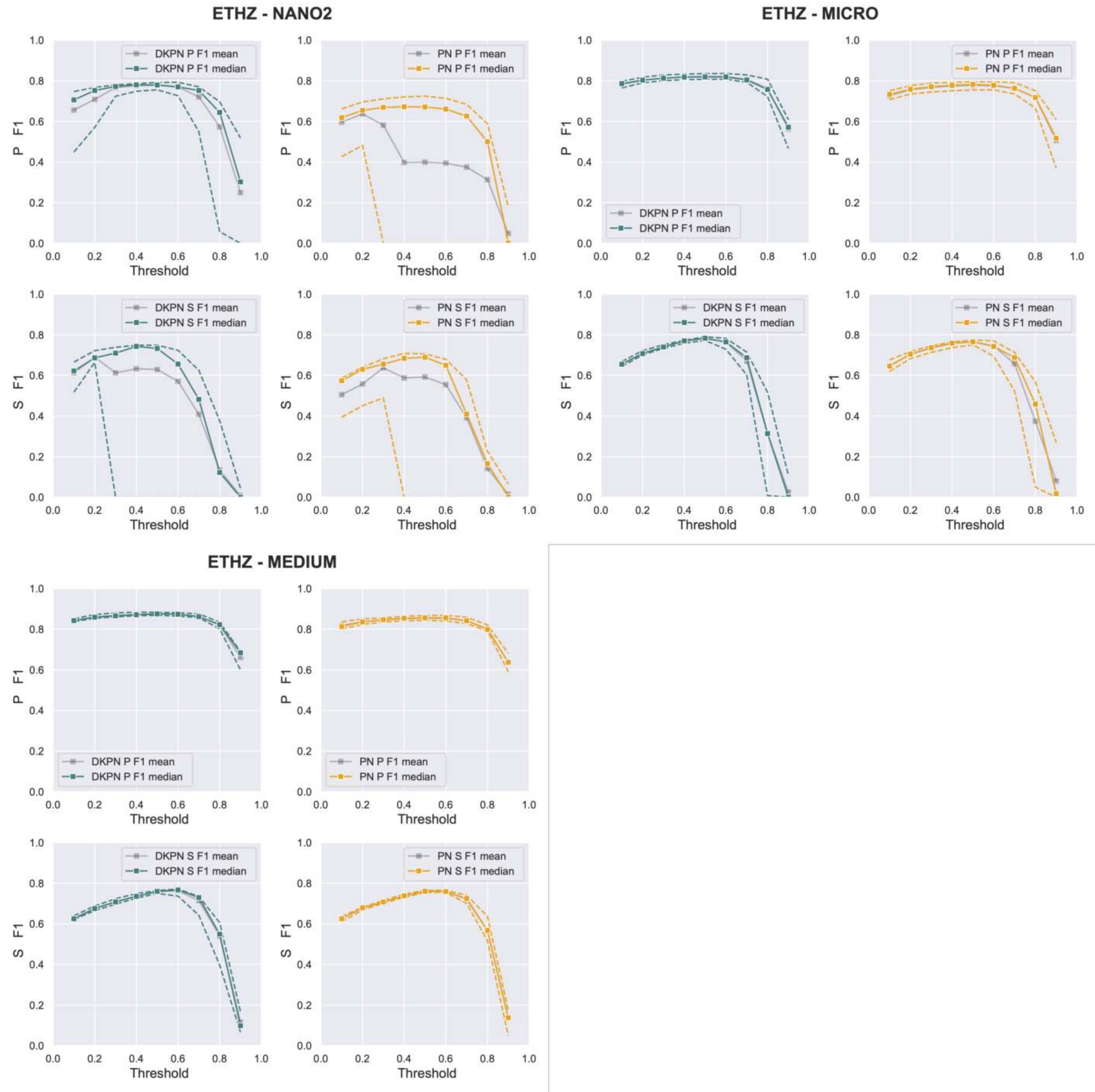


Figure 8 Comparison of PhaseNet (PN) and DKPN with testing on the ETHZ dataset (cross-domain). F1 score metrics across a range of pick amplitude thresholds for P arrivals and S arrivals for models trained with different size INSTANCE datasets. Mean, median and upper/lower limits of the mean (dashed curves) of F1 for 7 runs each with 5,000 evaluation samples drawn from the test dataset. Because some ETHZ traces have missing P or S picks, FP count may be overestimated.

different events and datasets can have a similar overall form and amplitude even when the corresponding raw seismograms have very different absolute amplitudes and frequency content.

3.2 Test 2 – Cross-domain—INSTANCE training and ETHZ test datasets

A most important, general and realistic use case is where a pre-trained picker model will be applied cross-domain—to seismogram waveforms with substantially different characteristics from the training waveforms. The differences may be related to recording instruments, data-loggers, available channel types and gain, wave propagation, site conditions and noise, and the distance, size, stress-drop and other properties of target seismic sources. To examine the cross-domain case, we apply models trained with the INSTANCE datasets of different sizes to testing (i.e. application) on waveforms from the ETHZ and PNW datasets. Sample waveforms are presented in Supplementary Files S2 and S3. The resulting F1 metrics across a range of pick amplitude thresholds for ETHZ test datasets are shown in Figure 8.

For P arrivals, relative to in-domain testing with INSTANCE (Test 1; Figure 6), DKPN shows a small reduction in overall performance and stability (e.g. maxi-

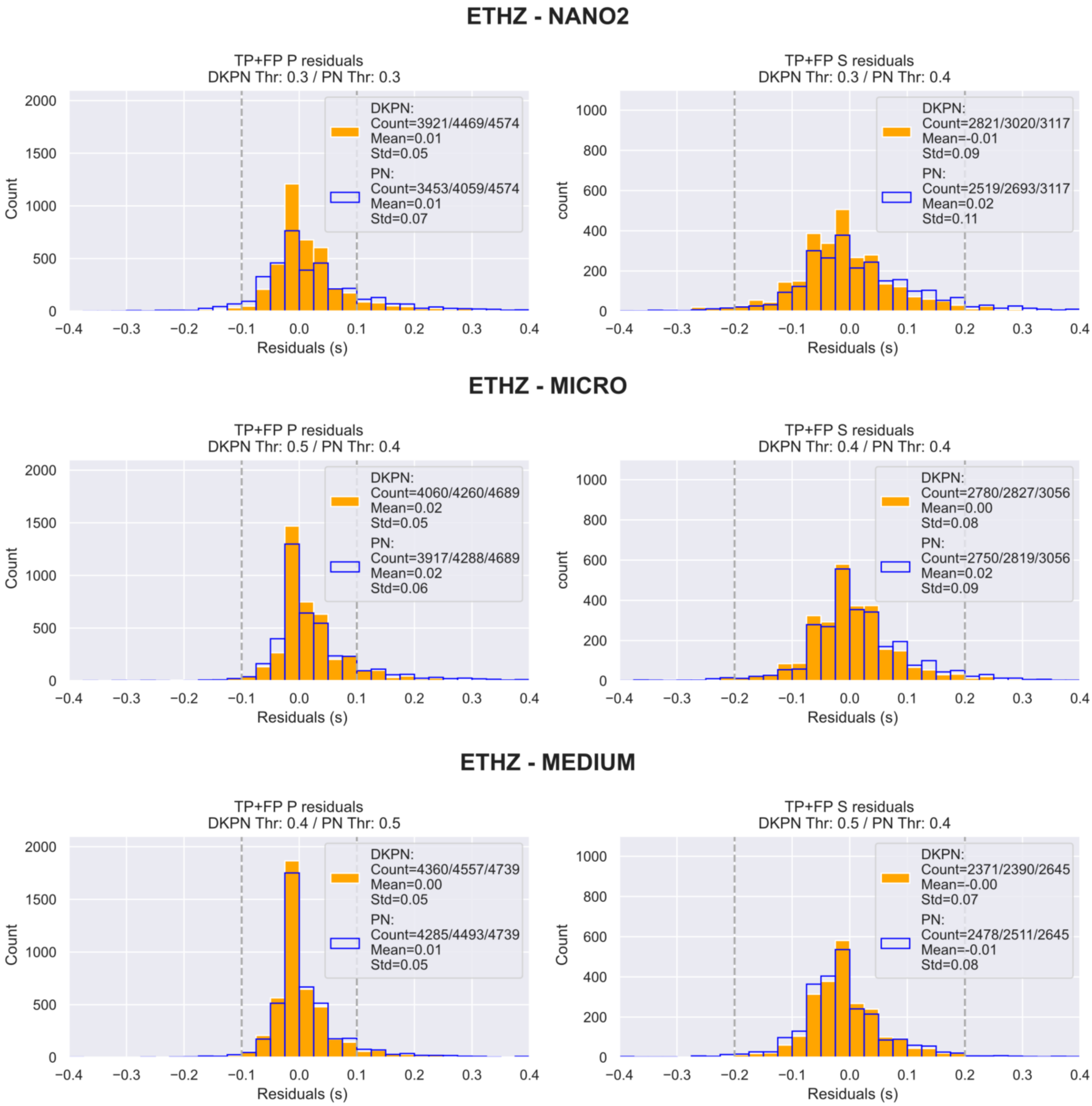


Figure 9 Histograms of P and S pick residuals for selected INSTANCE training dataset sizes tested on the ETHZ dataset (cross-domain). Results shown for the pick amplitude threshold (Thr) giving the highest F1 score for each case of dataset size and method (DKPN or PhaseNet). Vertical, dashed gray lines show the maximum difference between a pick time and the corresponding label time (0.1 sec for P and 0.2 sec for S) to declare a correctly predicted arrival for evaluation statistics. Pick counts show: Number of residuals (number of predicted picks) used in the mean, mode and standard-deviation statistics (trimmed within twice ΔTp: ± 0.2 sec for P and ± 0.4 sec for S) / Total number of residuals available / Total number of label picks available for the test case.

mum F1 scores almost always ≥ ~0.8 and converging to ~0.9 for larger training datasets) while PhaseNet shows a slightly larger reduction in performance (e.g. maximum F1 scores around 0.7 for the smaller datasets, and converging to ~0.85 for larger datasets). With the smaller training sets (NANO2 and MICRO) DKPN shows better results than PhaseNet, which indicates that, for application to ETHZ data, the DKPN model, with domain-knowledge input processing, has intrinsic properties that improve generalization and effective P picking with cross-domain data, as well as allowing use of smaller training datasets.

For S arrivals (Figure 8), as with Test 1, the performance of DKPN and PhaseNet are notably poorer than for P arrivals, though DKPN performs slightly better for the two smaller datasets. Relative to in-domain testing with INSTANCE (Figure 6), both DKPN and PhaseNet show generally reduced performance (e.g. maximum F1 scores of ~0.7-0.75 instead of ~0.8) except for an increase in Recall, due to a decrease in false negative

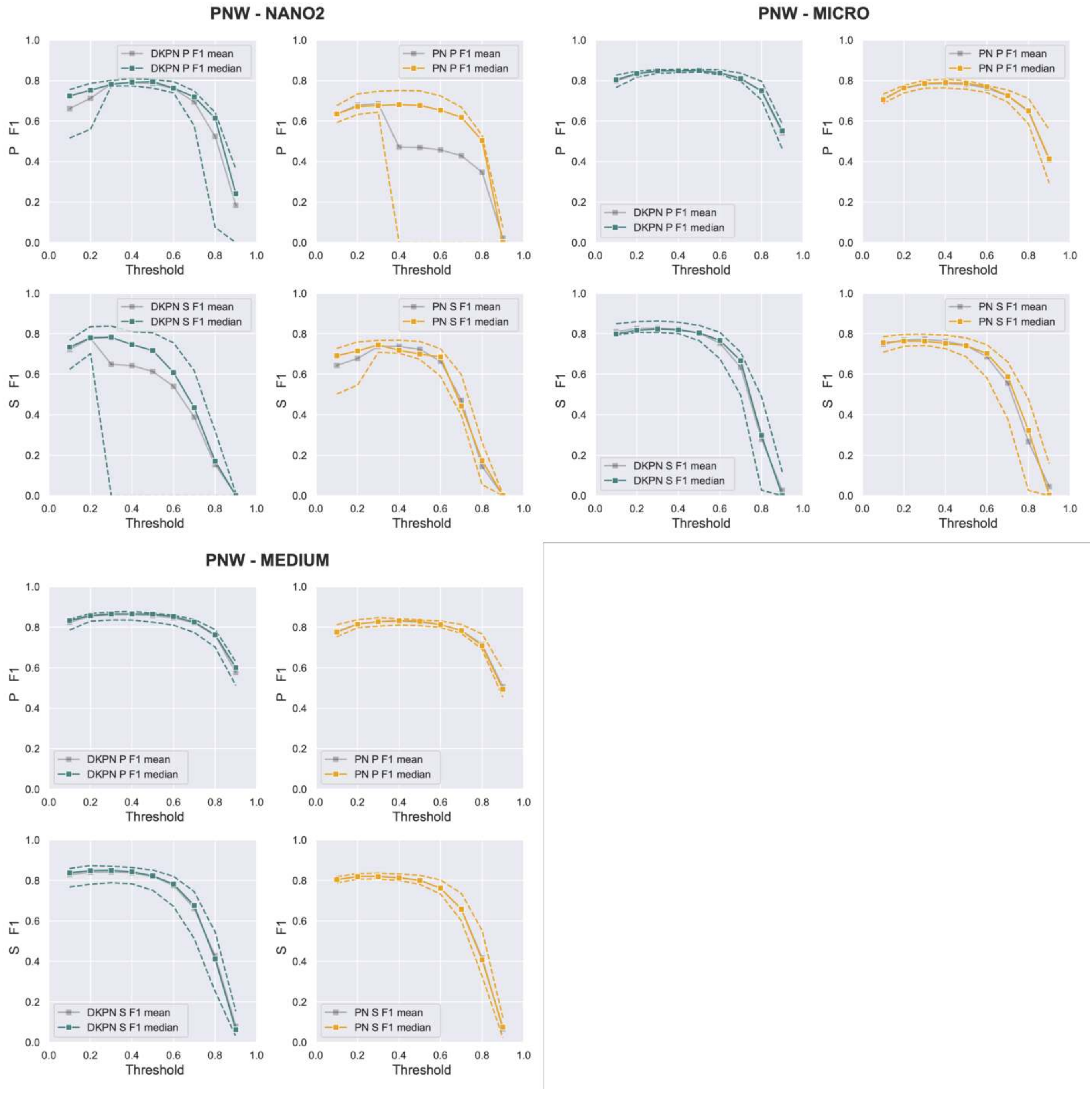


Figure 10 Comparison of PhaseNet (PN) and DKPN with testing on the PNW dataset (cross-domain). F1 score metrics across a range of pick amplitude thresholds for P arrivals and S arrivals for models trained with different size INSTANCE datasets. Mean, median and upper/lower limits of the mean (dashed curves) of F1 for 7 runs each with 5,000 evaluation samples drawn from the test dataset. Because some PNW traces have missing P or S picks, FP count may be overestimated.

count (Supplementary Text S2) .

Histograms of P and S pick residuals for PhaseNet and DKPN are shown in Figure 9. For P, in contrast to the INSTANCE testing results, the total number of predicted picks within twice ΔT_p increases with increasing training dataset size and is greater for DKPN than for PhaseNet. There is little change with respect to training dataset size in the mean of the residuals, which is always near zero, or for standard deviation. For P, DKPN always has a higher count of near-zero residual picks than PhaseNet, especially for the smallest training datasets (e.g. NANO2), in agreement with the evolution of F1 score and other statistics discussed above. For S, as with INSTANCE testing, DKPN has a higher count of near-zero residual picks than PhaseNet for the smaller training datasets, while for larger training datasets both models show almost identical performance. Notably, the total number of predicted picks within twice ΔT_p decreases (DKPN) or is roughly stable (PhaseNet) instead of increasing with increasing training dataset size as for P.

3.3 Test 3 – Cross-domain—INSTANCE training and PNW test datasets

We examine a second cross-domain case, applying the models trained with the INSTANCE datasets of different sizes to testing on waveforms from the PNW dataset. Relative to INSTANCE and ETHZ, the PNW waveform dataset is characterized by many trace sets with clear, impulsive S arrivals at larger S-P times (larger epicentral distance), but also trace sets with missing horizontal channels. Since the DKPN processing requires 3 component trace sets, we only use data for PNW which includes all 3 channels (Supplementary File S3). The resulting F1 metrics across a range of pick amplitude thresholds for PNW test datasets are shown in Figure 10.

For P arrivals, the PNW results are similar to those for cross-domain testing with ETHZ (Test 2; Figure 8) with a small reduction in overall performance and stability relative to in-domain testing with INSTANCE (Test 1; Figure 6) (e.g. maximum F1 scores ~0.8 instead of ~0.9 for larger datasets) and a small performance increase of DKPN over PhaseNet for larger training sets (MICRO and MEDIUM), and a more prominent increase for the smallest datasets (NANO2).

For S arrivals (Figure 10), as with ETHZ (Test 2; Figure 8), the performance of DKPN and PhaseNet for the smallest training dataset, NANO2 is slightly poorer than in-domain testing with INSTANCE (Test 1; Figure 6), nearly identical for the MICRO dataset, and, for the largest dataset, MEDIUM, nearly identical for PhaseNet and slightly improved for DKPN. These latter results are surprising for a cross-domain dataset, likely explained by the high rate in the PNW dataset of clear, impulsive S arrivals which may resemble well S arrivals captured most strongly in INSTANCE training, and, for DKPN, by the sensitivity of the introduced domain knowledge to impulsive arrivals.

Histograms of P and S pick residuals for PhaseNet and DKPN for PNW testing are shown in Figure 11. For P, and similar to cross-domain, ETHZ testing, with increasing training dataset size the total pick rate generally increases, there is little evolution for mean (always near zero) and standard-deviation, while DKPN has a higher count of near-zero residual picks than PhaseNet for all training dataset sizes. For S, DKPN shows slightly better statistics and count of near-zero residuals than PhaseNet in agreement with the evolution of F1 score and other S statistics discussed above. However, both PhaseNet and DKPN show a consistent negative mean residual of almost 0.1 sec, perhaps suggesting that the impulsiveness of many PNW S onsets relative to typical S onsets on INSTANCE training waveforms is leading the INSTANCE trained network to bias and advance the pick time predictions relative to those for INSTANCE waveforms.

4 Discussion

We use classical picker algorithms as domain-knowledge to transform raw seismogram waveforms into modified input features for the deep-learning PhaseNet picker, without otherwise modifying the picker architecture. We compare the deep-learning picker with modified input, DKPN, with standard PhaseNet when both are trained using the same INSTANCE data, training strategy and hyper-parameters and applied to an INSTANCE in-domain and two cross-domain datasets, ETHZ and PNW.

For P detection and picking, cross-domain application to the ETHZ (Test 2; Figure 8) and PNW datasets (Test 3; Figure 10) shows an improvement in F1 of around 15% for DKPN over PhaseNet for the smallest training dataset NANO2. For larger cross-domain training datasets and for all INSTANCE in-domain testing (Test 1; Figure 6) PhaseNet and DKPN show almost identical performance. These results suggest that with smaller size training datasets the domain-knowledge modified input of the DKPN deep-learning model provides useful prior information for effective and stable seismic phase detection and picking. The DKPN network thus does not need to learn this information during training (Figure 12), though the basic PhaseNet architecture is still capable of efficiently learning equivalent information during training with larger datasets. In histograms of P pick residuals (Figures 7, 9 and 11), DKPN generally shows a higher count of near-zero residual picks than PhaseNet, with slight reduction of this difference for the largest training datasets. This suggests the domain-knowledge modified input of the DKPN provides some improvement in the fine-scale onset timing of picks over PhaseNet. Overall, besides pick detection, much of the training for both PhaseNet and DKPN likely involves refinement of onset timing, phase identification and other pick characterization tasks; these are difficult tasks in manual tuning of classical picker algorithms and perhaps fundamentally better addressed with machine-learning optimization (Vassallo et al., 2012; Yeck et al., 2020).

The improved DKPN performance for P picking relative to PhaseNet for the ETHZ and PNW datasets with smaller training dataset sizes is likely due to the ETHZ and PNW pre-event noise, event waveforms, and P and S phase onsets having greater differences from the INSTANCE training waveforms than can be accommodated by the generalization of the PhaseNet INSTANCE training with small datasets. Important differences in waveforms relative to INSTANCE may include a larger number of regional events with lower frequency waveforms in ETHZ (Supplementary File S2), and the prevalence of sharper S onsets in PNW (Supplementary File S3).

These learning and performance differences indicate that for P picking, relative to purely data-driven deep-learning pickers such as PhaseNet, DKPN or other domain-knowledge machine-learning pickers may be better for small to very small training sets, may generalize better, e.g. when applied cross-domain to waveforms having very different characteristics to the waveforms of the training events, and may provide generally smaller differences relative to manual pick timing.

In addition, for all NANO2 models (Figure 6, 8, 10), DKPN is more stable than PhaseNet in P performances across many threshold levels as indicated by the spread of upper/lower limits of the mean (dashed curves). This

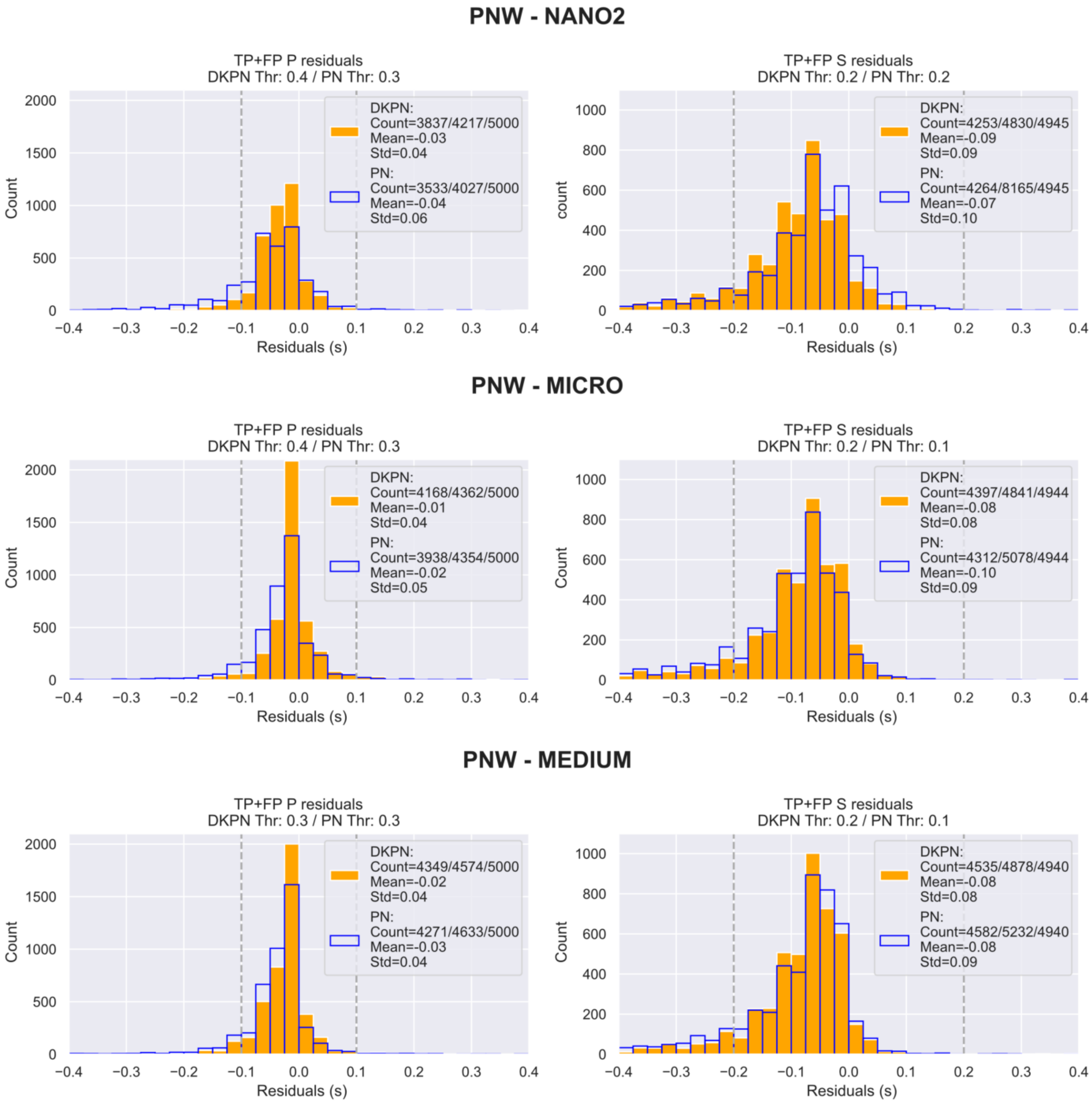


Figure 11 Histograms of P and S pick residuals for selected INSTANCE training dataset sizes tested on the ETHZ dataset (cross-domain). Results shown for the pick amplitude threshold (Thr) giving the highest F1 score for each case of dataset size and method (DKPN or PhaseNet). Vertical, dashed gray lines show the maximum difference between a pick time and the corresponding label time (0.1 sec for P and 0.2 sec for S) to declare a correctly predicted arrival for evaluation statistics. Pick counts show: Number of residuals (number of predicted picks) used in the mean, mode and standard-deviation statistics (trimmed within twice ΔTp: ± 0.2 sec for P and ± 0.4 sec for S) / Total number of residuals available / Total number of label picks available for the test case.

means that DKPN is less sensitive to changes in pick threshold selection, proving to be more assertive about onset prediction (i.e. sharper prediction probability functions) even when few training data are available. DKPN is also less sensitive to training-data selection as resulting from different random selections across the 7 training-testing experiments, as shown from the upper and lower bounds of F1-scores that better follow the median trends.

For S detection and picking, both PhaseNet and DKPN show lower performance relative to P detection and picking (Figures 6-11), as also found for the deep-learning pickers examined in (Münchmeyer et al., 2022). This reduced performance is most likely due to S arrivals occurring in the P coda, and to the often emergent and complicated form of S arrival onsets, especially at regional distances in areas of complex geology. For in-domain testing on INSTANCE datasets and cross-domain ETHZ testing, PhaseNet and DKPN show nearly identical S performance for all but the smallest training datasets. However, for the cross-domain PNW test dataset DKPN shows slightly better S picking performance than PhaseNet across all training dataset sizes. This result may be related to the PNW dataset, relative to INSTANCE and ETHZ, having a large number of clear, impulsive S arrivals, which may match well impulsive P arrivals for which classical pickers such as Filter Picker are optimized, and thus more easily detected by DKPN.

Training dataset size is an important issue with deep-learning pickers, as there are few large, well curated and error-free seismic waveform datasets with reliable, manual or other, reference picks. Many studies, such as temporary aftershock monitoring and short-term experiments, may have manually picked datasets that are too small for training with pure, data-driven picker models. Moreover, waveforms for some studies may have specific characteristics (e.g., in frequency content, epicentral distance ranges, noise, distribution of magnitudes) that preclude processing with machine learning methods pretrained with large datasets with different waveform characteristics. Here we have used relatively small to moderate size training datasets (~800 to 245k samples) relative to other key studies (e.g. 11k, 65k, 555k, 780k, 1.3M and 4.5M training and evaluation traces for the 6 picker models examined in Münchmeyer et al., 2022). We have shown that DKPN sometimes outperforms PhaseNet with smaller datasets for P picking, especially for cross-domain picking of the ETHZ and PNW datasets, probably due to the prior, domain-knowledge information on picking inherent in the DKPN input traces (Figure 12). Thus domain-knowledge based methods such as DKPN may be especially useful for studies with smaller datasets, especially those with unusual waveform characteristics which necessitates picker retraining, as well as for when limited computing time or resources are available. The combination of domain-knowledge based methods with transfer learning (e.g., Jozinović et al., 2021) may be particularly useful with small datasets that require retraining of machine learning pickers.

The FilterPicker CF amplifies and transforms energy onsets and changes in frequency content into abrupt, step- or pulse-like waveforms, while remaining fairly insensitive to absolute amplitudes and frequency content which vary between events and datasets (Figs 2-5). Improvements in P picking performance of DKPN over the purely data-driven PhaseNet may primarily be due to the similarity between these CF waveforms and the narrow, Gaussian wavelets of the target, probabilistic, picks (Fig. 12). To help verify this proposition, we ran a version of DKPN which retains the 3 channels of raw waveform input, giving 8 channels total for input. This change gives almost no difference in the picking results such as mean and median F1 scores, except for a degradation of results for S picking with the smallest training dataset NANO2, and the 8 channel input leads to increased spread of the upper/lower limits of the mean for the NANO2 and MICRO training datasets. This reduction in performance suggest the 5 channels of CF's plus inclination and modulus waveforms input to DKPN retain the majority of information relevant to picking effectively contained in the raw waveforms.

The DKPN network thus apparently receives rule-based, *deterministically* modified input that resembles a simple transformation of the required output defining pick detection, potentially simplifying training and improving performance and stability, and also providing an inherent mechanism for generalization. On the other hand, given a classical picker CF, the design and optimization of subsequent algorithms for refining onset-timing, phase identification and other characterization are difficult and somewhat *haphazard* tasks (Lomax et al., 2012; Vassallo et al., 2012). In DKPN and PhaseNet these subsequent tasks are performed by the deep-neural-network; indeed, high-dimensional, *stochastically*-driven machine-learning is eminently suited to such tasks. However, when observations from a network of seismometers are available, a domain-knowledge, rule-based approach may also be valuable for pick characterization tasks such as quality control (Ning et al., 2022). And, in practice, domain-knowledge is used to improve detection and picking even with nominally, data-driven, machine-learning pickers, since many studies apply a high-pass filter to suppress known microseismic noise at longer period and amplify arrivals of interest expected at higher frequencies (Mousavi et al., 2019, 2020; Münchmeyer et al., 2022; Ross et al., 2018b,a; Woollam et al., 2019).

Additional study might investigate the usage of "simpler" and "shallower" model-architectures than that of PhaseNet, while still feeding the DKPN input or similar. Such a configuration could help understand the effects of domain-knowledge on machine-learning model generalization. In particular, less complex architectures may allow easier setting of meta-parameters during the learning stages and better explanation of the machine learning models, and produce more robust models that are easier to debug and improve. However, if pick characterization tasks other than detection account for much of the learning effort during training for both DKPN and PhaseNet, then the use of CF, inclination and modulus waveforms in DKPN is not likely to allow reducing the number of layers or otherwise simplifying the underlying CNN architecture inherited from PhaseNet.

FilterPicker and STA/LTA methods in general require stabilization after the start of a time-series and after impulsive arrivals; the time of stabilization for FilterPicker is proportional to the long-term, recursive-filter time-averaging scale. This stabilization, besides making it necessary to have sufficient background data before the first arrival in a time-series, usually degrades picker sensitivity to arrivals closely following previous arrivals, in particular an S arrival, even when higher amplitude than the preceding P. This is one reason why we include in DKPN the instantaneous polarization modulus time-series, which preserves S amplitude relative to P, and the inclination time-series, which often changes character from predominantly up-down to near horizontal at the S arrival. Future work might investigate if, in the context of a domain-knowledge, machine-learning picker, it is possible to modify the FilterPicker CF algorithm to reduce adverse effects of the stabilization without otherwise adversely affecting the overall picker per-

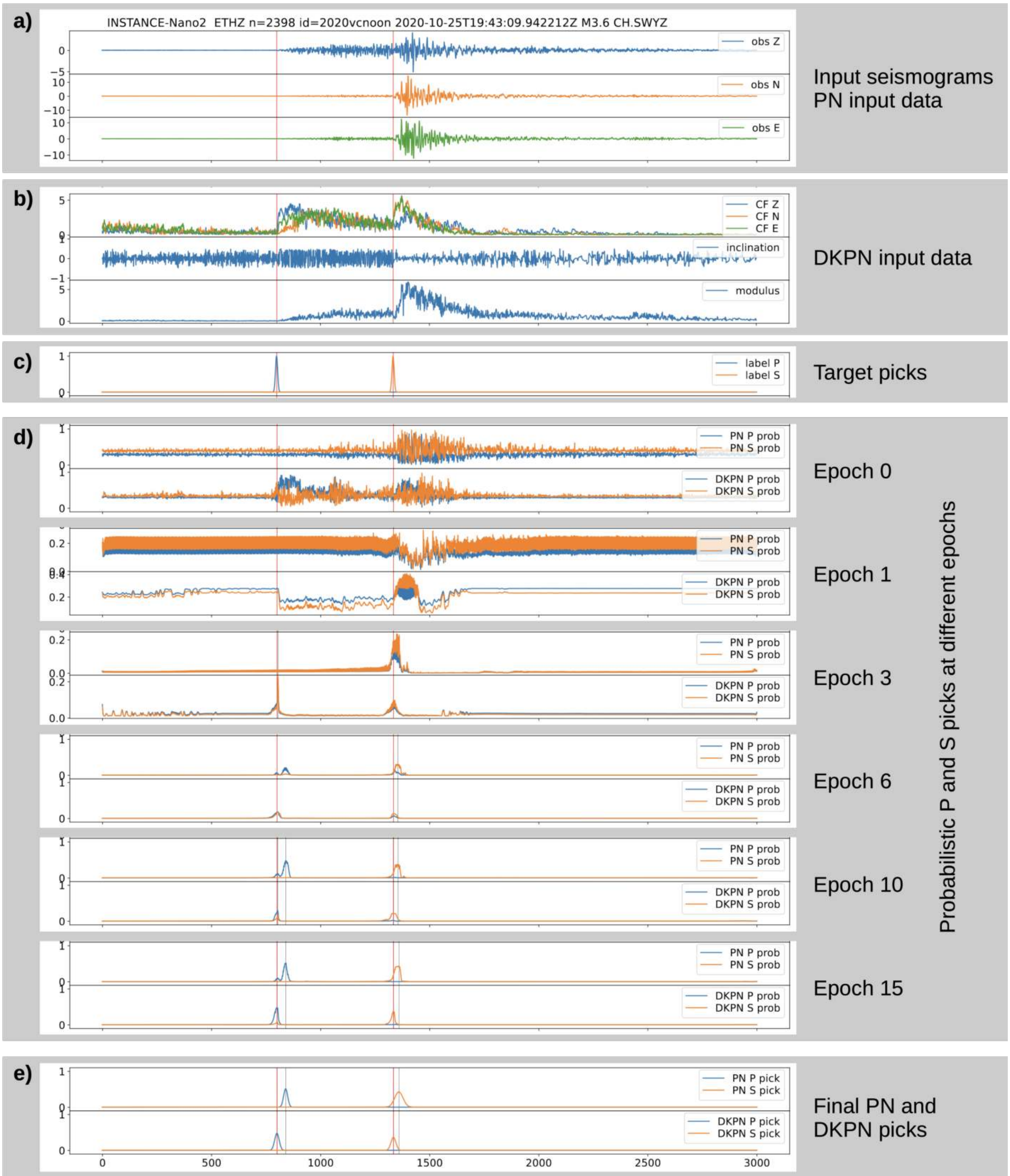


Figure 12 Results with progression of epoch for models trained with the INSTANCE NANO2 dataset applied to ETHZ trace samples. Panel a) shows de-meaned and normalized Z, N, E component input seismograms; panel b) normalized, DKPN FilterPicker Z, N, E CFs, and normalized instantaneous inclination and modulus; and panel c) P (blue) and S (orange) pick label data (red vertical lines). Panel e) shows PhaseNet (PN) and DKPN Gaussian P (blue) and S (orange) predicted picks after epoch 15 training. Panel d) shows PhaseNet and DKPN probabilistic P (blue) and S (orange) pick predictions for a sequence of training epochs; for clarity, the predictions for epochs 1 and 3 are not normalized. For the untrained model (epoch 0) the predictions are random, non-linear mappings of the input traces. The epoch 0 predictions for DKPN reflect the input CFs and polarization traces and show a distinct P arrival signature which is not present in the predictions for PhaseNet, which reflect mainly amplitudes in the near-raw seismograms (continued).

Figure 12 (Continued) The DKPN probabilistic predictions show in epoch 1 the P arrival as a step-like signal and the S arrival as a concentrated prediction, in epoch 3 both P and S arrivals as isolated predictions, and from epoch 6 to 15 as stable predictions, though in epoch 6 the P arrival has both P and S predictions. PhaseNet obtains an S but not P prediction in epoch 3, isolated but noisy P and S predictions starting from epoch 6, and stable predictions between epochs 10 and 15. The slower evolution of the PhaseNet predictions from near random to clear arrivals through epochs 0-6 support that it is learning both arrival detection and picking throughout the training process. P predictions at the S arrival time for both PhaseNet and DKPN, and S predictions at the P arrival for DKPN visible in epochs 3 and 6 are highly suppressed through learning by epoch 10. The final PhaseNet P and S picks are delayed relative to the label picks, likely due to the emergent amplitude of the arrivals which is not represented well in the INSTANCE training dataset. The final DKPN picks do not show this delay, likely due to the high sensitivity of the domain-knowledge preprocessing (panel b) to changes in waveform characteristics besides amplitude, such as frequency content.

formance.

Relative to PhaseNet, DKPN has an increase in overall training and evaluation time due to the preprocessing required to derive the 3-component, FilterPicker CFs, modulus and inclination from the seismogram waveforms. However, our calculations in this study show that the processing time penalty is effectively removed through code optimization and use of parallel, GPU processing. In addition, preprocessing the training dataset once before training and storing it on disk or in memory can remove most of the DKPN training time penalty. Moreover, despite having an increased dimensionality of input data, DKPN training converges faster (requires fewer epochs) than PhaseNet (see Supplementary File S4).

5 Conclusions

Using seismological domain-knowledge, we transform 3-component seismograms into the characteristic functions of a classical, multi-band picker, plus instantaneous modulus and inclination. We replace the near-raw seismogram input of the deep-learning picker PhaseNet with these transformed traces, forming DKPN, a modified PhaseNet, and we compare the performance of DKPN and standard PhaseNet with different training and testing datasets. DKPN shows some improvements in performance and generalization over PhaseNet, and may be applicable with smaller training datasets. DKPN requires more computation time than standard PhaseNet due to the additional, domain-knowledge preprocessing. However, this time penalty can be removed with code optimization, GPU use, real-time processing, and storing DKPN preprocessed waveforms for training. Additionally, DKPN training time (number of epochs) may be reduced relative to PhaseNet.

For P arrivals, DKPN shows little or no improvement in performance over PhaseNet in picking the in-domain, INSTANCE dataset for all training dataset sizes, and in picking cross-domain ETHZ and PNW datasets for larger training dataset sizes. These results demonstrate the power and robustness of the PhaseNet architecture for extracting information relevant to pick detection and characterization from near-raw seismogram waveforms. However, DKPN generally shows improved statistics such as true positive rate and increased number of picks with small residuals, and sometimes improved metrics such as F1 score, especially for small training datasets and for cross-domain testing. These improvements can be attributed to the additional information relevant to picking introduced in the DKPN input data by the rule-based, domain-knowledge waveform preprocessing. For the purely data-driven PhaseNet, much of this same picking-specific information must be learned by the network in training; the efficiency and success of this training will depend on the training dataset being sufficiently large and having similar event waveform characteristics to the application datasets.

The performance of both PhaseNet and DKPN is worse for picking S arrivals than for P, likely mainly due to S onsets occurring in the P coda. Both models show similar S performance for in- and cross-domain picking on the ETHZ dataset, but DKPN shows slightly better performance than PhaseNet for cross-domain S picking on the PNW dataset, likely due to the frequent occurrence of sharp S onsets on the PNW waveforms which are less prevalent in the INSTANCE training data.

Overall, our results show that PhaseNet, and perhaps deep-neural-network pickers in general, have a sufficiently large and complex architecture to learn to accurately map key characteristics of seismogram waveforms and phase onset energy into detections and picks, including for cross-domain application. This learning requires comprehensive and high-quality, but not necessarily very large training datasets.

However, given our results, DKPN is of interest for cross-domain picking when retraining on the target dataset is not possible, or for cases where training is needed but can be performed on only a very small dataset, such as when few manual picks are available. Further work with DKPN and other, domain-knowledge augmented machine-learning procedures for picking and other seismological analyses is warranted to investigate performance improvements over pure, data-driven, learning algorithms, especially for small or highly varied training datasets and for strongly cross-domain application.

Acknowledgments

We gratefully thank Mostafa Mousavi and an anonymous reviewer for extensive and helpful comments, and the editor Yen Joe Tan and all Seismica volunteers for their help and contributions. This research was supported by the SOME (Seismological Oriented Machine lEarning) project - INGV Pianeta Dinamico 2021 Tema 8 (grant no. CUP D53J1900017001) funded by the Italian Ministry of University and Research "Fondo finalizzato al rilancio degli investimenti delle amministrazioni centrali dello Stato e allo sviluppo del Paese, legge 145/2018". M. Bagagli was partially supported by the European Union's research and innovation program under the Marie Sklodowska Curie Action, grant agreement (101105516).

6 Data and code availability

All datasets used in this study are available in SeisBench (https://www.github.com/seisbench/seisbench). Python codes used for training, testing and visualization presented in Tests 1, 2 and 3, and run with PyTorch (https://pytorch.org) and SeisBench (https://www.github.com/seisbench/seisbench), are available at https://github.com/INGV/DKPN.

The supplementary information for this article includes a pdf file containing additional text, table and figures describing the experiment setup, training and testing stages, Supplementary Files S1-S3 contain example waveforms and PhaseNet and DKPN processing for the three datasets, and Supplementary File S4 contain train-validation loss-curves for selected training runs.

All calculations were performed on a GAIA multi GPU server (equipped with 4 80GB NVidia A100) running the ICE4AI software stack by E4 Analytics (https://www.e4company.com/en/gpu-appliance-for-artificial-intelligence) and providing environments with PyTorch (https://pytorch.org), SeisBench (https://www.github.com/seisbench/seisbench), and ObsPy (Beyreuther et al., 2010; Krischer et al., 2015, , http://obspy.org). Word processing and some figures were done with LibreOffice (https://www.libreoffice.org).

7 Competing interests

The authors declare no conflicts of interest with respect to the research, authorship, and publication of this article.

References

Akazawa, T. A technique for automatic detection of onset time of P-and S-phases in strong motion records. In *Proc. of the 13th World Conf. on Earthquake Engineering*, Vancouver, Canada, 2004. http://www.iitk.ac.in/nicee/wcee/article/13_786.pdf.

Allen, R. Automatic phase pickers: Their present use and future prospects. *Bulletin of the Seismological Society of America*, 72 (6B):S225–S242, Dec. 1982. doi: 10.1785/bssa07206b0225.

Allen, R. V. Automatic earthquake recognition and timing from single traces. *Bulletin of the Seismological Society of America*, 68(5): 1521–1532, Oct. 1978. doi: 10.1785/bssa0680051521.

Alvarez, I., Garcia, L., Mota, S., Cortes, G., Benitez, C., and De la Torre, A. An Automatic P-Phase Picking Algorithm Based on Adaptive Multiband Processing. *IEEE Geoscience and Remote Sensing Letters*, 10(6):1488–1492, Nov. 2013. doi: 10.1109/lgrs.2013.2260720.

Anant, K. S. and Dowla, F. U. Wavelet transform methods for phase identification in three-component seismograms. *Bulletin of the Seismological Society of America*, 87(6):1598–1612, Dec. 1997. doi: 10.1785/bssa0870061598.

Baer, M. and Kradolfer, U. An automatic phase picker for local and teleseismic events. *Bulletin of the Seismological Society of America*, 77(4):1437–1445, Aug. 1987. doi: 10.1785/bssa0770041437.

Bagagli, M. *Seismicity and seismic tomography across scales: application to the greater Alpine region*. PhD thesis, 2022. doi: 10.3929/ETHZ-B-000580361.

Bai, C.-y. Automatic Phase-Detection and Identification by Full Use of a Single Three-Component Broadband Seismogram. *Bulletin of the Seismological Society of America*, 90(1):187–198, Feb. 2000. doi: 10.1785/0119990070.

Balestriero, R. and Baraniuk, R. Mad Max: Affine Spline Insights into Deep Learning, 2018. doi: 10.48550/ARXIV.1805.06576.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., and Wassermann, J. ObsPy: A Python Toolbox for Seismology. *Seismological Research Letters*, 81(3):530–533, May 2010. doi: 10.1785/gssrl.81.3.530.

Beyreuther, M., Hammer, C., Wassermann, J., Ohrnberger, M., and Megies, T. Constructing a Hidden Markov Model based earthquake detector: application to induced seismicity: Constructing a HMM based earthquake detector. *Geophysical Journal International*, 189(1):602–610, Feb. 2012. doi: 10.1111/j.1365-246x.2012.05361.x.

Borghesi, A., Baldo, F., and Milano, M. Improving Deep Learning Models via Constraint-Based Domain Knowledge: a Brief Survey, 2020. doi: 10.48550/ARXIV.2005.10691.

Chen, C. and Holland, A. A. PhasePApy: A Robust Pure Python Package for Automatic Identification of Seismic Phases. *Seismological Research Letters*, 87(6):1384–1396, Aug. 2016. doi: 10.1785/0220160019.

Dai, H. and MacBeth, C. Automatic picking of seismic arrivals in local earthquake data using an artificial neural network. *Geophysical Journal International*, 120(3):758–774, Mar. 1995. doi: 10.1111/j.1365-246x.1995.tb01851.x.

Enescu, N. Seismic Data Processing Using Nonlinear Prediction and Neural networks. In *IEEE NORSIG Symposium*, Espoo, Finland, 1996.

Gentili, S. and Michelini, A. Automatic picking of P and S phases using a neural tree. *Journal of Seismology*, 10(1):39–63, Jan. 2006. doi: 10.1007/s10950-006-2296-6.

Hien, D. A guide to receptive field arithmetic for Convolutional Neural Networks, 2018. https://blog.mlreview.com/a-guide-to-receptive-field-arithmetic-for-convolutional-neural-networks-e0f514068807. Retrieved August 16, 2022, from.

Jozinović, D., Lomax, A., Štajduhar, I., and Michelini, A. Transfer learning: improving neural network based prediction of earthquake ground shaking for an area with insufficient training data. *Geophysical Journal International*, 229(1):704–718, Dec. 2021. doi: 10.1093/gji/ggab488.

Kim, A., Nakamura, Y., Yukutake, Y., Uematsu, H., and Abe, Y. Development of a high-performance seismic phase picker using deep learning in the Hakone volcanic area. *Earth, Planets and Space*, 75(1), May 2023. doi: 10.1186/s40623-023-01840-5.

Kong, Q., Trugman, D. T., Ross, Z. E., Bianco, M. J., Meade, B. J., and Gerstoft, P. Machine Learning in Seismology: Turning Data into Insights. *Seismological Research Letters*, 90(1):3–14, Nov. 2018. doi: 10.1785/0220180259.

Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C., and Wassermann, J. ObsPy: a bridge for seismology into the scientific Python ecosystem. *Computational Sci-*

ence & Discovery, 8(1):014003, May 2015. doi: 10.1088/1749-4699/8/1/014003.

LeCun, Y., Bengio, Y., and Hinton, G. Deep learning. *Nature*, 521 (7553):436–444, May 2015. doi: 10.1038/nature14539.

Liao, W.-Y., Lee, E.-J., Mu, D., Chen, P., and Rau, R.-J. ARRU Phase Picker: Attention Recurrent-Residual U-Net for Picking Seismic P- and S-Phase Arrivals. *Seismological Research Letters*, 92(4): 2410–2428, Mar. 2021. doi: 10.1785/0220200382.

Lomax, A., Satriano, C., and Vassallo, M. Automatic Picker Developments and Optimization: FilterPicker–a Robust, Broadband Picker for Real-Time Seismic Monitoring and Earthquake Early Warning. *Seismological Research Letters*, 83(3):531–540, May 2012. doi: 10.1785/gssrl.83.3.531.

Lomax, A., Michelini, A., and Curtis, A. *Earthquake Location, Direct, Global-Search Methods*, page 1–33. Springer New York, 2014. doi: 10.1007/978-3-642-27737-5_150-2.

Marcus, G. Innateness, AlphaZero, and Artificial Intelligence, 2018. doi: 10.48550/ARXIV.1801.05667.

McEvilly, T. V. and Majer, E. L. ASP: An Automated Seismic Processor for microearthquake networks. *Bulletin of the Seismological Society of America*, 72(1):303–325, Feb. 1982. doi: 10.1785/bssa0720010303.

Michelini, A., Cianetti, S., Gaviano, S., Giunchi, C., Jozinović, D., and Lauciani, V. INSTANCE – the Italian seismic dataset for machine learning. *Earth System Science Data*, 13(12):5509–5544, Nov. 2021. doi: 10.5194/essd-13-5509-2021.

Mousavi, S. M. and Beroza, G. C. Deep-learning seismology. *Science*, 377(6607), Aug. 2022. doi: 10.1126/science.abm4470.

Mousavi, S. M., Langston, C. A., and Horton, S. P. Automatic microseismic denoising and onset detection using the synchrosqueezed continuous wavelet transform. *GEOPHYSICS*, 81 (4):V341–V355, July 2016. doi: 10.1190/geo2015-0598.1.

Mousavi, S. M., Zhu, W., Sheng, Y., and Beroza, G. C. CRED: A Deep Residual Network of Convolutional and Recurrent Units for Earthquake Signal Detection. *Scientific Reports*, 9(1), July 2019. doi: 10.1038/s41598-019-45748-1.

Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., and Beroza, G. C. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature Communications*, 11(1), Aug. 2020. doi: 10.1038/s41467-020-17591-w.

Mousset, E., Cansi, Y., Crusem, R., and Souchet, Y. A connectionist approach for automatic labeling of regional seismic phases using a single vertical component seismogram. *Geophysical Research Letters*, 23(6):681–684, Mar. 1996. doi: 10.1029/95gl03811.

Muralidhar, N., Islam, M. R., Marwah, M., Karpatne, A., and Ramakrishnan, N. Incorporating Prior Domain Knowledge into Deep Neural Networks. In *2018 IEEE International Conference on Big Data (Big Data)*. IEEE, Dec. 2018. doi: 10.1109/bigdata.2018.8621955.

Münchmeyer, J., Woollam, J., Rietbrock, A., Tilmann, F., Lange, D., Bornstein, T., Diehl, T., Giunchi, C., Haslinger, F., Jozinović, D., Michelini, A., Saul, J., and Soto, H. Which Picker Fits My Data? A Quantitative Evaluation of Deep Learning Based Seismic Pickers. *Journal of Geophysical Research: Solid Earth*, 127(1), Jan. 2022. doi: 10.1029/2021jb023499.

Ni, Y., Hutko, A., Skene, F., Denolle, M., Malone, S., Bodin, P., Hartog, R., and Wright, A. Curated Pacific Northwest AI-ready Seismic Dataset. *Seismica*, 2(1), May 2023. doi: 10.26443/seismica.v2i1.368.

Ning, I. L. C., Swafford, L., Craven, M., Davies, K., Earnest, E., and Thornton, D. Automation of passive seismic processing via machine learning and physics-informed methods. In *Second International Meeting for Applied Geoscience & Energy*. Society of Exploration Geophysicists and American Association of Petroleum Geologists, Aug. 2022. doi: 10.1190/image2022-3750116.1.

Njirjak, M., Otović, E., Jozinović, D., Lerga, J., Mauša, G., Michelini, A., and Štajduhar, I. The Choice of Time–Frequency Representations of Non-Stationary Signals Affects Machine Learning Model Accuracy: A Case Study on Earthquake Detection from LEN-DB Data. *Mathematics*, 10(6):965, Mar. 2022. doi: 10.3390/math10060965.

Plešinger, A., Hellweg, M., and Seidl, D. Interactive high-resolution polarization analysis of broad-band seismograms. *Journal of Geophysics*, 59(1):129–139, 1986. https://journal.geophysicsjournal.com/JofG/article/view/203.

Ronneberger, O., Fischer, P., and Brox, T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*, page 234–241. Springer International Publishing, 2015. doi: 10.1007/978-3-319-24574-4_28.

Ross, Z. E. and Ben-Zion, Y. Automatic picking of direct P, S seismic phases and fault zone head waves. *Geophysical Journal International*, 199(1):368–381, Aug. 2014. doi: 10.1093/gji/ggu267.

Ross, Z. E., Meier, M., and Hauksson, E. P Wave Arrival Picking and First-Motion Polarity Determination With Deep Learning. *Journal of Geophysical Research: Solid Earth*, 123(6):5120–5129, June 2018a. doi: 10.1029/2017jb015251.

Ross, Z. E., Meier, M., Hauksson, E., and Heaton, T. H. Generalized Seismic Phase Detection with Deep Learning. *Bulletin of the Seismological Society of America*, 108(5A):2894–2901, Aug. 2018b. doi: 10.1785/0120180080.

Satriano, C., Lomax, A., and Zollo, A. Real-Time Evolutionary Earthquake Location for Seismic Early Warning. *Bulletin of the Seismological Society of America*, 98(3):1482–1494, June 2008. doi: 10.1785/0120060159.

Sleeman, R. and van Eck, T. Robust automatic P-phase picking: an on-line implementation in the analysis of broadband seismogram recordings. *Physics of the Earth and Planetary Interiors*, 113(1–4):265–275, June 1999. doi: 10.1016/s0031-9201(99)00007-2.

Soto, H. and Schurr, B. DeepPhasePick: A method for detecting and picking seismic phases from local earthquakes based on highly optimized convolutional and recurrent deep neural networks. *Geophysical Journal International*, July 2021. doi: 10.1093/gji/ggab266.

Stevenson, P. R. Microearthquakes at Flathead Lake, Montana: A study using automatic earthquake processing. *Bulletin of the Seismological Society of America*, 66(1):61–80, Feb. 1976. doi: 10.1785/bssa0660010061.

Vassallo, M., Satriano, C., and Lomax, A. Automatic Picker Developments and Optimization: A Strategy for Improving the Performances of Automatic Phase Pickers. *Seismological Research Letters*, 83(3):541–554, May 2012. doi: 10.1785/gssrl.83.3.541.

Vidale, J. Complex polarization analysis of particle motion. *Bulletin of the Seismological Society of America*, 76(5):1393–1405, 1986. doi: 10.1785/BSSA0760051393.

Wang, J. and Teng, T.-L. Artificial neural network-based seismic detector. *Bulletin of the Seismological Society of America*, 85(1): 308–319, Feb. 1995. doi: 10.1785/bssa0850010308.

Withers, M., Aster, R., Young, C., Beiriger, J., Harris, M., Moore, S., and Trujillo, J. A comparison of select trigger algorithms for automated global seismic phase and event detection. *Bulletin of the Seismological Society of America*, 88(1):95–106, Feb. 1998. doi: 10.1785/bssa0880010095.

Woollam, J., Rietbrock, A., Bueno, A., and De Angelis, S. Convolutional Neural Network for Seismic Phase Classification, Performance Demonstration over a Local Seismic Network. *Seismological Research Letters*, 90(2A):491–502, Jan. 2019. doi: 10.1785/0220180312.

Woollam, J., Münchmeyer, J., Tilmann, F., Rietbrock, A., Lange, D., Bornstein, T., Diehl, T., Giunchi, C., Haslinger, F., Jozinović, D., Michelini, A., Saul, J., and Soto, H. SeisBench—A Toolbox for Machine Learning in Seismology. *Seismological Research Letters*, 93(3):1695–1709, Mar. 2022. doi: 10.1785/0220210324.

Yeck, W. L., Patton, J. M., Ross, Z. E., Hayes, G. P., Guy, M. R., Ambruz, N. B., Shelly, D. R., Benz, H. M., and Earle, P. S. Leveraging Deep Learning in Global 24/7 Real-Time Earthquake Monitoring at the National Earthquake Information Center. *Seismological Research Letters*, 92(1):469–480, Sept. 2020. doi: 10.1785/0220200178.

Yu, Z. and Wang, W. LPPN: A Lightweight Network for Fast Phase Picking. *Seismological Research Letters*, 93(5):2834–2846, June 2022. doi: 10.1785/0220210309.

Zhang, H., Thurber, C., and Rowe, C. Automatic P-Wave Arrival Detection and Picking with Multiscale Wavelet Analysis for Single-Component Recordings. *Bulletin of the Seismological Society of America*, 93(5):1904–1912, Oct. 2003. doi: 10.1785/0120020241.

Zhu, W. and Beroza, G. C. PhaseNet: A Deep-Neural-Network-Based Seismic Arrival Time Picking Method. *Geophysical Journal International*, Oct. 2018. doi: 10.1093/gji/ggy423.

doi:10.26443/seismica.v3i1.1157

SEISMICA

Apparent Non-Double-Couple Components as Artifacts of Moment Tensor Inversion

Boris Rösler * [1,2], **Seth Stein** [1,3], **Adam T. Ringler** [4], **Jiří Vackář** [5]

[1]Department of Earth and Planetary Sciences, Northwestern University, Evanston, Illinois, U.S.A., [2]now at Center for Scientific Research and Higher Education at Ensenada (CICESE), Ensenada, Mexico, [3]Institute for Policy Research, Northwestern University, Evanston, Illinois, U.S.A., [4]Albuquerque Seismological Laboratory, United States Geological Survey, Albuquerque, New Mexico, U.S.A., [5]Institute of Rock Structure and Mechanics, Czech Academy of Sciences, Prague, Czech Republic

Author contributions: *Conceptualization*: Boris Rösler. *Methodology*: Boris Rösler, Seth Stein. *Software*: Boris Rösler, Adam T. Ringler, Jiří Vackář. *Validation*: Seth Stein. *Formal Analysis*: Boris Rösler, Seth Stein. *Writing - Original draft*: Boris Rösler. *Writing - Review & Editing*: Seth Stein, Adam T. Ringler, Jiří Vackář.

Abstract Compilations of earthquake moment tensors from global and regional catalogs find pervasive non-double-couple (NDC) components with a mean deviation from a double-couple (DC) source of around 20%. Their distributions vary only slightly with magnitude, faulting mechanism, or geologic environments. This consistency suggests that for most earthquakes, especially smaller ones whose rupture processes are expected to be simpler, the NDC components are largely artifacts of the moment tensor inversion procedure. This possibility is also supported by the fact that NDC components for individual earthquakes with $M_w < 6.5$ are only weakly correlated between catalogs. We explore this possibility by generating synthetic seismograms for the double-couple components of earthquakes around the world using one Earth model and inverting them with a different Earth model. To match the waveforms with a different Earth model, the inversion changes the mechanisms to include a substantial NDC component while largely preserving the fault geometry (DC component). The resulting NDC components have a size and distribution similar to those reported for the earthquakes in the Global Centroid Moment Tensor (GCMT) catalog. The fact that numerical experiments replicate general features of the pervasive NDC components reported in moment tensor catalogs implies that these components are largely artifacts of the inversions not adequately accounting for the effects of laterally varying Earth structure.

Production Editor:
Gareth Funning
Handling Editor:
Stephen Hicks
Copy & Layout Editor:
Ethan Williams

Received:
23 November 2023
Accepted:
5 March 2024
Published:
4 April 2024

1 Introduction

Moment tensors (MTs) are a general description of earthquake sources, providing information beyond a double-couple (DC) force system representing slip on a fault plane (Gilbert, 1971). The deployment of global digital seismic networks allowed development of large catalogs of moment tensors (e.g., Ekström et al. 2012). These catalogs have become an important tool in studies worldwide, including analyzing global plate motions and deformation in plate boundary zones and within plates.

As MT catalogs were developed (Dziewonski et al., 1981), it became clear that many earthquakes showed non-double-couple (NDC) components whose origin became a topic of investigation (Sipkin, 1986; Frohlich, 1994). This effect is evident along many plate boundaries (Fig. 1). For example, along the Mid-Atlantic ridge, many earthquake mechanisms deviate from the DC mechanisms for pure strike-slip on transforms and normal faulting on ridge segments (Tréhu et al., 1981).

A NDC component is identified by decomposing a moment tensor. Diagonalization of the MT yields a tensor with eigenvalues λ_1, λ_2 and λ_3 on its diagonal, where $\lambda_1 > \lambda_3 > \lambda_2$. Subtracting a diagonal matrix with components equal to the isotropic moment $M_0^{\text{iso}} = (\lambda_1 + \lambda_2 + \lambda_3)/3$, representing the source's volumetric change, yields the deviatoric moment tensor typically reported in catalogs. The deviatoric MT has no net volume change because its trace, the sum of its eigenvalues, $\lambda_1' + \lambda_2' + \lambda_3' = 0$.

The decomposition of this deviatoric MT into a DC and a NDC component (Knopoff and Randall, 1970) describes the NDC component as a compensated linear vector dipole (CLVD), three force dipoles with one twice the magnitude of the others, yielding no volume change,

$$\begin{pmatrix} \lambda_1' & 0 & 0 \\ 0 & \lambda_2' & 0 \\ 0 & 0 & \lambda_3' \end{pmatrix} = (\lambda_1' + 2\lambda_3') \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \lambda_3' \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (1)$$

For a pure double couple, $\lambda_3' = 0$ and $\lambda_1' = -\lambda_2'$. The ratio of the smallest and absolutely largest eigenvalues

$$|\epsilon| = \frac{|\lambda_3'|}{\max(|\lambda_1'|, |\lambda_2'|)} \quad (2)$$

quantifies the size of the NDC component, the deviation from a DC source (Dziewonski et al., 1981).

NDC components can arise in several ways. Some appear to reflect intrinsically complex source processes

*Corresponding author: boris@earth.northwestern.edu

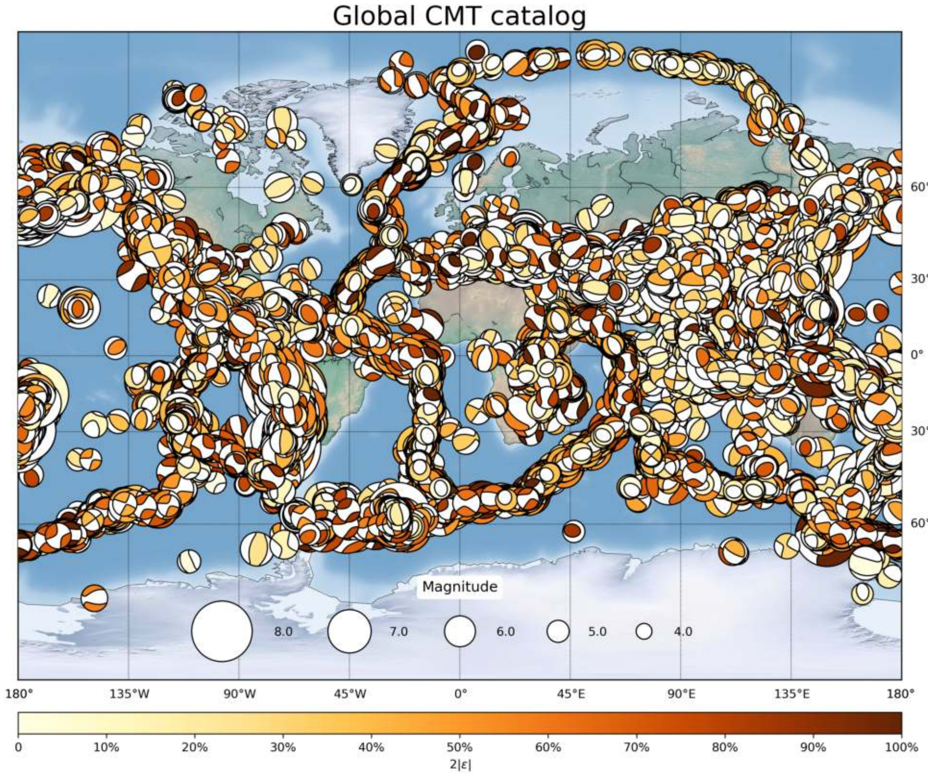


Figure 1 The source mechanisms of the 62337 earthquakes in the Global Centroid Moment Tensor (GCMT) catalog between 1976 and 2022. Many mechanisms show a substantial non-double couple (NDC) component.

that differ from slip on a planar fault for earthquakes in specific geologic environments, notably volcanic areas (e.g., Kanamori and Given, 1982; Ross et al., 1996; Nettles and Ekström, 1998; Shuler et al., 2013a,b; Gudmundsson et al., 2016; Sandanbata et al., 2021; Rodríguez-Cardozo et al., 2021). Others are additive, reflecting the combined effect of near-simultaneous rupture on multiple faults with different geometries (e.g., Kawakatsu, 1991; Hayes et al., 2010; Hamling et al., 2017; Scognamiglio et al., 2018; Yang et al., 2021; Ruhl et al., 2021), or a rupture with changes in geometry (Wald and Heaton, 1994; Cohee and Beroza, 1994; Pang et al., 2020). Alternatively, they may be artifactual (McNamara et al., 2013; Ammon et al., 1994), results of the inversion without geologic meaning, generated by noise in the waveforms (Šílený et al., 1996; Jechumtálová and Šílený, 2001), inappropriate seismic station coverage (Cesca et al., 2006; Ford et al., 2010; Vera Rodriguez et al., 2011; Domingues et al., 2013), or not accounting for laterally varying Earth structure during the inversion (Šílený, 2004; Cesca et al., 2006; Rößler et al., 2007).

By comparing MTs in the Global CMT Project (GCMT) and U.S. Geological Survey (USGS) catalogs, Rösler et al. (2021) found that the distribution (e.g., mean and standard deviation) of NDC components in different catalogs are quite similar. However, the actual values are only weakly correlated for events with $M_w < 6.5$. Because the catalogs use different inversion procedures, the poor correlation between the NDC components suggests that they are artifacts. Moreover, using a large dataset compiled from multiple global and regional MT catalogs, Rösler and Stein (2022) found that for earthquakes with magnitudes $2.9 < M_w < 8.2$, NDC components are common, with an average value of $2|\epsilon| = 23.2\%$ that varies only slightly with magnitude. They argued that this consistency indicates that NDC components are unlikely to reflect rupture on multiple faults, which is more likely to occur for large earthquakes (Quigley et al., 2017). They also found only small differences in NDC components between earthquakes with different faulting mechanisms, or in different geologic environments. These results are interesting in that NDC components are often assumed to be most likely in volcanic and thus extensional environments (Ross et al., 1996; Julian and Sipkin, 1985; Miller et al., 1998; Nettles and Ekström, 1998). Hence the consistency suggests that for most earthquakes, especially smaller ones, the NDC components do not reflect complex rupture processes and are therefore artifacts of the moment tensor inversion.

Studies have identified the influence of Earth structure on moment tensor inversions (Henry et al., 2002; Hjörleifsdóttir and Ekström, 2010), and attempts have been made to reduce the uncertainties introduced by lateral heterogeneity in the Earth in inversions based on one-dimensional (1D) Earth models (Vasyura-Bathke et al., 2021; Phạm and Tkalčić, 2021). Apart, moment tensor solutions based on Green's functions generated for three-dimensional (3D) Earth models for regional earthquakes have been calculated (Hingee et al., 2011; Hejrani et al., 2017; Zhu and Zhou, 2016; Wang and Zhan, 2020; Liu et al., 2004; Jechumtálová and Bulant, 2014; Covellone and Savage, 2012), which, in some cases, were improved using rotational motions (Donner et al., 2020). Sawade et al. (2022) compiled the CMT3D moment tensor catalog for global earthquakes with M_w

≥ 5.5 using a 3D Earth model for the inversion. Its NDC components are smaller on average than the ones in the GCMT catalog, showing how spurious NDC components can be reduced in inversions accounting for lateral variations in the structure of the Earth.

The availability of multiple moment tensor catalogs with different inversion procedures allows us to quantify this phenomenon on a global scale, and we explore whether the distribution of NDC components can be generated by the MT inversion process by generating synthetic seismograms for pure DC source processes and invert them using an approach that simulates the effect of uncertainties in the Earth structure assumed for the inversion.

2 Methodology

For our study, we model earthquakes selected from the Global Centroid Moment Tensor (GCMT) catalog, which contains about 60,000 earthquakes since 1976 (Ekström et al., 2012). We classify earthquakes by their faulting type, following Frohlich (1992). We calculate the plunge of the P- (most compressive), N- (null), and T-axes (least compressive) from the eigenvectors of the moment tensors. An earthquake is considered a normal faulting earthquake if its P-axis plunge satisfies $\sin^2 \delta_P \geq 2/3$ ($\delta_P \geq 54.75°$), strike-slip if its N-axis plunge exceeds $54.75°$, and a thrust fault if its T-axis plunge exceeds $54.75°$ (Saloor and Okal, 2018). If the plunge of none of the axes exceeds the threshold, an earthquake is considered oblique faulting. Of the earthquakes in the catalog up to the year 2022, 34.9% have thrust mechanisms, 24.5% are strike-slip, 22.2% have normal faulting mechanisms, and 18.4% have oblique-faulting mechanisms. We therefore model nine earthquakes with thrust mechanisms, six with strike-slip mechanisms, five with normal faulting mechanisms, and five with oblique faulting mechanism, consistent with the fractions in the GCMT catalog. We choose these 25 earthquakes as representative of the geologic environment they occurred in, with a geographical distribution over all continents to avoid bias due to Earth's elliptical shape and its rotation (Fig. 2). 22 of them have depths of less than 30km, similar to the vast majority of earthquakes in the GCMT catalog. To avoid magnitude bias, we model the earthquakes with a moment magnitude of M_w 7.0, which ensures that they are detected at stations around the world, but can be modeled as a point source.

The one-dimensional Preliminary Reference Earth model (PREM, Dziewonski and Anderson, 1981) is used by the three global MT catalogs of the Global CMT Project (Dziewonski et al., 1981; Ekström et al., 2012), the USGS (Hayes et al., 2009) and the Deutsches GeoForschungsZentrum (GFZ, Joachim Saul, personal communication, 2022) in their inversion procedure. To simulate the deviation of the Earth model used in the inversion from the actual Earth structure, we perturb the one-dimensional model for the generation of synthetic seismograms. We calculate synthetic seismograms by way of normal mode summation for a spherically symmetric non-rotating Earth and include the effects of attenuation and self-gravitation for both the spheroidal and toroidal components from angular order 0 to 8000 with a period range of 0 to 20 mHz and a sampling rate of 1/s. The first 200 dispersion branches are also included in the synthetic seismogram calculation, and the moment tensor is incorporated by convolving the eigenfunctions with the moment tensor components obtained for the DC component from the MTs as reported in the GCMT catalog. We then invert them using Green's functions generated for the unperturbed PREM model. Any NDC component in the resulting moment tensors is thus an artifact of the inversion resulting from the difference in Earth structure. This process simulates current methodology for global catalogs in which MTs are found using one-dimensional (laterally homogeneous) Earth models. Although ideally the analysis would involve generating and inverting seismograms for three-dimensional (laterally varying) Earth models, our approach using a range of perturbations allows quantifying the effect of laterally varying Earth structure in the inversion.

We perturb the elastic and anelastic structures of the Earth model independently while retaining the elastic moduli K, the bulk modulus, and μ, the shear modulus (Dahlen and Tromp, 1998, Section 3.6.2) in each layer from the PREM model. We determine μ and K from the PREM P- and S-wave velocities v_{pv} and v_{sv}, and density ρ using

$$\begin{aligned} v_{sv} &= \sqrt{\frac{K}{\mu}} \Rightarrow \mu = v_{sv}^2 \rho \\ v_{pv} &= \sqrt{\frac{K + \frac{4}{3}\mu}{\rho}} \Rightarrow K = \rho \left(v_{pv}^2 - \tfrac{4}{3} v_{sv}^2 \right) . \end{aligned} \tag{3}$$

Next, we perturb the P-wave velocity v_{pv} by randomly choosing its value from a Gaussian distribution with mean v_{pv} and standard deviation 1, 3, 5, and 10%, which reflect the uncertainties in the velocities (Dalton and Ekström, 2006). Keeping K and μ constant, we then determine the new density in each layer as

$$\rho_{new} = \frac{K + \frac{4}{3}\mu}{v_{pv,\ new}^2} \tag{4}$$

from which we determine the new S-wave velocity as

$$v_{sv} = \sqrt{\frac{\mu}{\rho_{new}}}. \tag{5}$$

Of the 185 layers in the model PREM, nine in the upper mantle are anisotropic. For those, η (Dahlen and Tromp, 1998, Section 8.9) which relates the horizontal P-wave velocity to the vertical S-wave velocity is given,

$$\eta = \frac{F}{\rho_{old} \left(v_{ph,\ old}^2 - 2 v_{sv,\ old}^2 \right)}, \tag{6}$$

from which we obtain $F = \eta \rho_{old} \left(v_{ph,\ old}^2 - 2 v_{sv,\ old}^2 \right)$. The new horizontal P-wave velocity is then

$$v_{ph,\ new} = \sqrt{\frac{F}{\eta \rho_{new}} + 2 v_{sv,\ new}^2}. \tag{7}$$

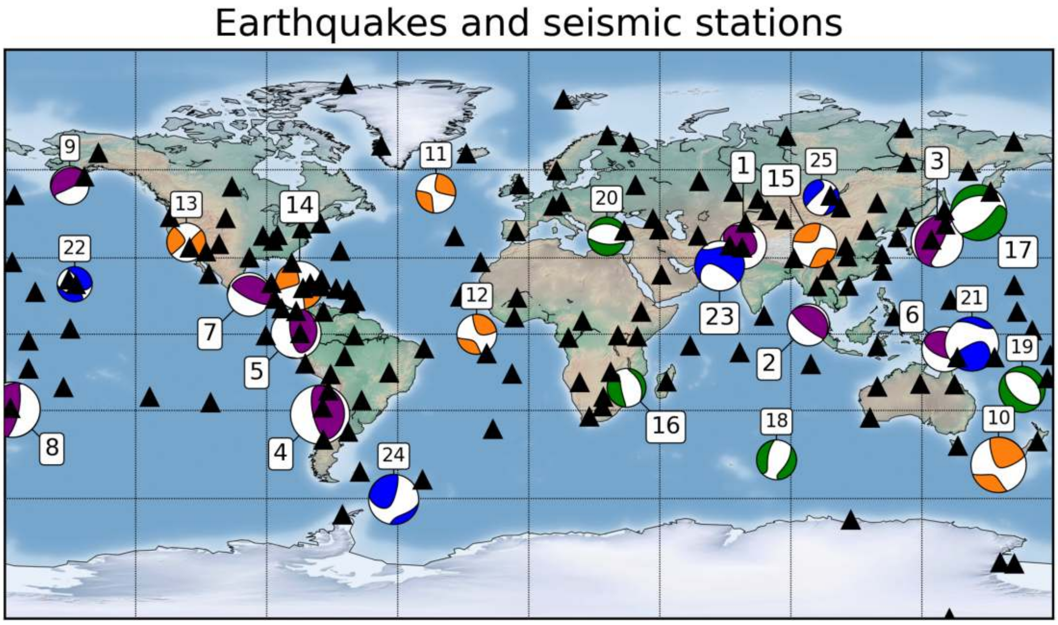


Figure 2 Earthquakes and seismic stations used in this study. The location and origin time of the earthquakes are listed in Table 1. Consistent with the fractions in the Global Centroid Moment Tensor (GCMT) catalog, nine of the 25 earthquakes have a thrust faulting mechanism, six have a strike-slip faulting mechanism, five have a normal faulting mechanism, and 5 have an oblique faulting mechanism. The 150 seismic stations in the Global Seismographic Network (GSN) network provide good worldwide station coverage.

Number	Date	Time	Latitude	Longitude	Depth	M_w	Faulting Type
1	2005-10-08	03:50:40.8	34.54	73.59	12.0	7.58	thrust
2	2010-05-09	05:59:41.6	3.75	96.02	37.2	7.25	thrust
3	2011-03-11	06:15:45.0	36.13	140.23	29.0	7.89	thrust
4	2015-09-16	22:54:32.9	-31.57	-71.67	17.4	8.27	thrust
5	2016-04-16	23:58:36.9	0.35	-79.93	22.3	7.78	thrust
6	2018-02-25	17:44:44.1	-6.07	142.75	12.0	7.47	thrust
7	2020-06-23	15:29:04.3	15.88	-96.01	21.5	7.38	thrust
8	2021-03-04	19:28:33.2	-29.72	-177.28	33.9	8.07	thrust
9	2021-08-14	11:57:43.5	55.18	-157.64	25.0	7.00	thrust
10	2004-12-23	14:59:04.4	-49.31	161.35	27.5	8.08	strike-slip
11	2015-02-13	18:59:12.2	52.65	-31.9	25.2	7.07	strike-slip
12	2016-08-29	04:29:57.9	-0.05	-17.83	26.8	7.10	strike-slip
13	2019-07-06	03:19:53.0	35.77	-117.6	12.0	7.03	strike-slip
14	2020-01-28	19:10:24.9	19.42	-78.76	23.9	7.69	strike-slip
15	2021-05-21	18:04:13.6	34.59	98.24	12.0	7.42	strike-slip
16	2006-02-22	22:19:07.8	-21.32	33.58	12.0	7.01	normal
17	2007-01-13	04:23:21.2	46.24	154.52	12.0	8.10	normal
18	2015-12-04	22:25:00.1	-47.62	85.09	28.9	7.10	normal
19	2018-12-05	04:18:08.4	-21.95	169.43	17.8	7.53	normal
20	2020-10-30	11:51:27.4	37.91	26.78	12.0	6.99	normal
21	2000-11-16	04:54:56.7	-3.98	152.17	24.0	8.00	oblique
22	2006-10-15	17:07:49.2	19.88	-155.93	48.0	6.71	oblique
23	2013-09-24	11:29:48.0	26.97	65.52	12.0	7.70	oblique
24	2013-11-17	09:04:55.5	-60.27	-46.4	23.8	7.78	oblique
25	2021-01-11	21:32:59.0	51.28	100.44	13.9	6.79	oblique

Table 1 List of earthquakes used in this study as reported in the GCMT catalog, indicating their origin time, location, depths, magnitude, and faulting type.

The new horizontal S-wave velocity v_{sh} is not restricted by the elastic constants and can thus be perturbed independently, similarly to the vertical P-wave velocity v_{pv}. Because the values related to anelastic structure are much less understood than the elastic structure (Karaoğlu and Romanowicz, 2018), the anelastic structure is perturbed by randomly choosing a value from Gaussian distributions whose means, $1/Q_\kappa$ and $1/Q_\mu$, are the inverse of the original P- and S- wave quality factors, and whose standard deviations of 25, 50, and 75% thereof are typical for the variations found in attenuation tomography studies (e.g., Dalton and Ekström,

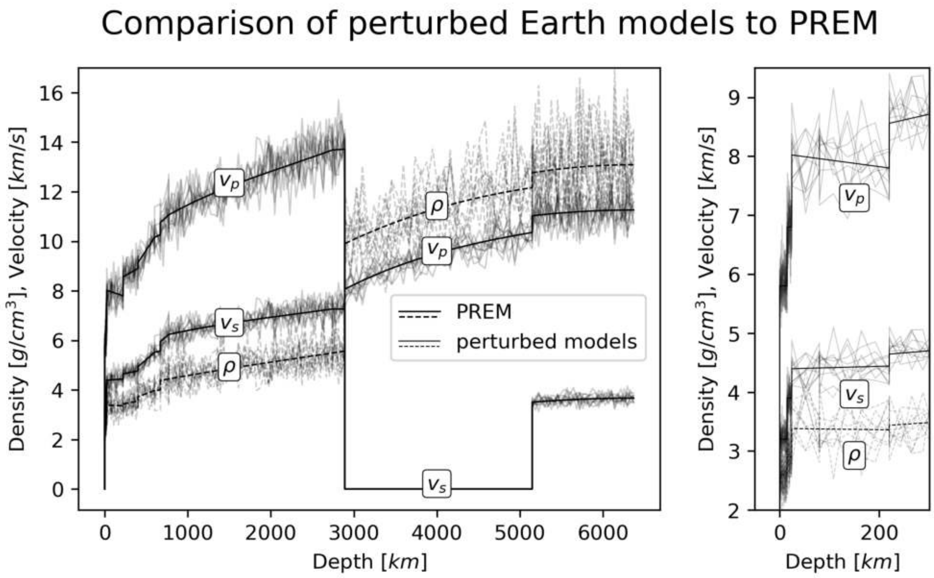


Figure 3 Comparison of the elastic structure of the Earth models used in this study, based on perturbation of model Preliminary Reference Earth Model (PREM)

2006).

Using these perturbed Earth models (Fig. 3), we generate three-component seismograms for the 150 stations of the Global Seismographic Network (GSN, Ringler et al., 2022) from one hour before the event time until 8000 s after it. This guarantees that the surface waves are included in the synthetic seismograms for all stations. For each earthquake we use the fifty closest stations which have an epicentral distance of at least 10° to avoid near-source effects (Aki and Richards, 2002).

To invert the seismograms, we use Green's functions generated for the unperturbed PREM model and perform a least-squares centroid MT inversion in a Bayesian framework using BayesISOLA (Vackář et al., 2017). BayesISOLA finds the moment tensor elements that give the best-fitting match to the synthetic seismograms in a full-waveform inversion from a linear combination of the Green's functions with the moment tensor elements as coefficients. We use frequencies from 0.002 to 0.01 Hz (i.e., periods between 100 and 500 s), as commonly used for global moment tensor inversions of large earthquakes (Ekström et al., 2012; Duputel et al., 2018; Kanamori and Rivera, 1993; Kanamori, 2008; Hayes et al., 2009), and the covariance matrix of Green's functions (Hallo and Gallovič, 2016) to estimate the effects of uncertainty in the Earth model. We fix the centroid location to the point for which the synthetic seismograms are generated but invert for centroid time in the range of ± 15 s about the centroid time of the synthetic event. We do not allow station-specific time delays as used in the inversion procedures of the GCMT catalog (Dziewonski et al., 1984). BayesISOLA calculates the best solution for the moment tensor as well as the posterior probability density function describing the uncertainty of the MT elements. Although the uncertainty might be useful to distinguish whether resulting NDC components are real or artifacts of the inversion, we only use the best solution in the comparison of the results with GCMT catalog, similar to the reported MTs in this catalog.

Similar to the procedure of the GCMT catalog, we constrain the isotropic component of the MTs during the inversion so that the resulting MTs are purely deviatoric ($\lambda_1' + \lambda_2' + \lambda_3' = 0$). However, in contrast to the GCMT procedure, we use longer period seismic waves. GCMT uses surface waves between 50 s and 150 s and mantle waves between 125 s and 350 s (Ekström et al., 2012). We do not use surface-waves delay dispersion maps, but the inversion procedures are comparable in most of parameters.

3 Results

Figure 4 shows an example of the seismic waveforms generated for the DC component of earthquake 12 in figure 2 at station II.SUR in Sutherland, South Africa, for a perturbed Earth model. Inversion of these seismograms using the unperturbed model PREM results in a moment tensor with an 8.7% artifactual NDC component, such that the waveforms generated for the resulting MT are nearly indistinguishable from those generated for the DC MT and the perturbed Earth model. Thus inversion with a different Earth model changes the mechanism to best match the waveforms at all stations by introducing a substantial NDC component. Repeating the process twelve times using different perturbed models to generate synthetic seismograms and inverting them with the unperturbed model yields a range of focal mechanisms (Fig. 5) with varying artifactual NDC components, averaging 4.4% for a perturbation of 5% in the elastic structure.

The DC component of the MT — the fault geometry — is generally retrieved well, as measured by the angle Φ in space required to rotate one set of a moment tensor's principal axes into the ones of another (Kagan, 1991). The angles are generally small, with an average value of 7.8°, less than the differences between moment tensors of the GCMT and the USGS for earthquakes of this mag-

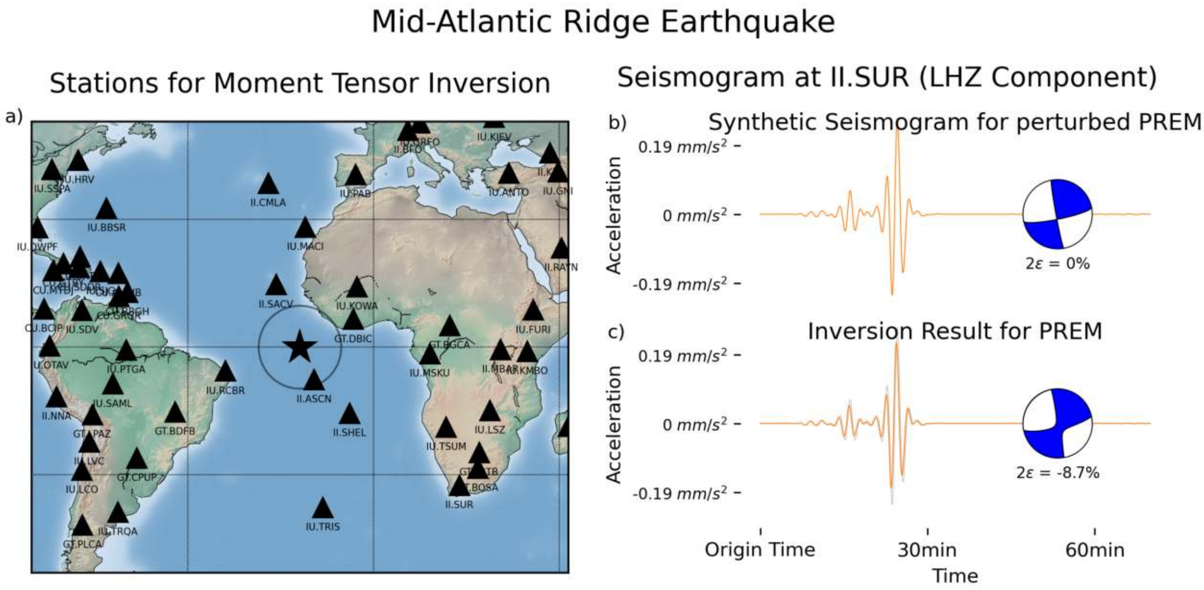


Figure 4 a) Stations for the moment tensor inversion of synthetic seismograms generated for the mid-Atlantic ridge earthquake of March 14, 1994. We used the fifty closest stations of the GSN network with an epicentral distance of at least $10°$. b) Synthetic seismogram at station II.SUR (Sutherland, South Africa) generated for the DC component of the moment tensor for a perturbed model PREM with standard deviation of 5% in the seismic velocities with periods of 100 to 500 s. c) Synthetic seismogram generated for the moment tensor resulting from the inversion performed using Green's functions generated for the unperturbed model PREM, compared to the input synthetic seismogram in b). Matching the waveform produces a spurious NDC component.

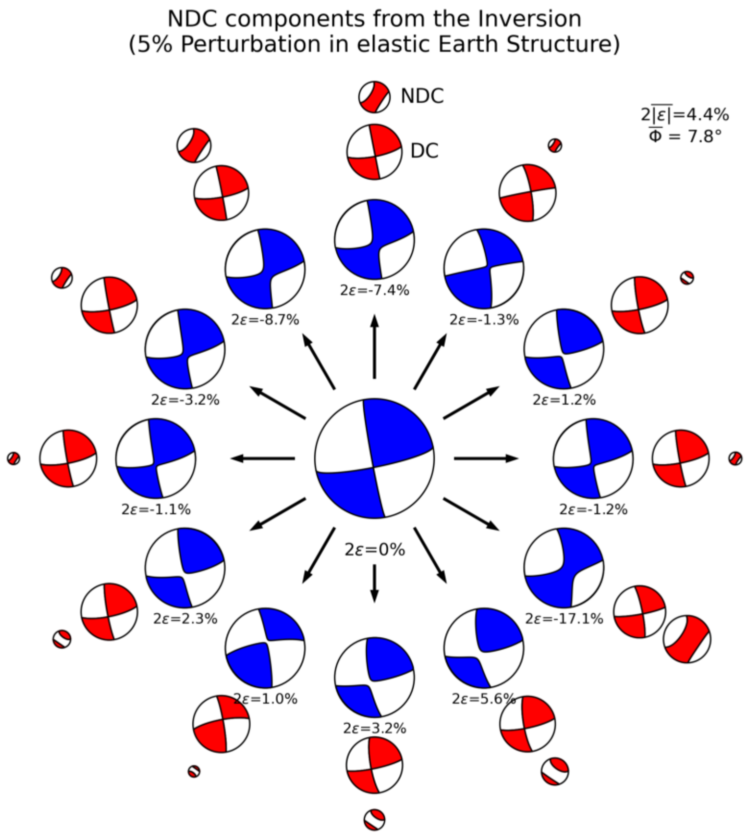


Figure 5 Inversion results for the mid-Atlantic ridge earthquake of March 14, 1994. Synthetic seismograms were generated for a pure double-couple (DC) mechanism and twelve different perturbed models based on Preliminary Reference Earth Model (PREM) and inverted using the unperturbed model. The resulting moment tensors (MTs) have very similar DC components, but substantial non-double couple (NDC) components (shown as twice as large) that differ in polarity and size.

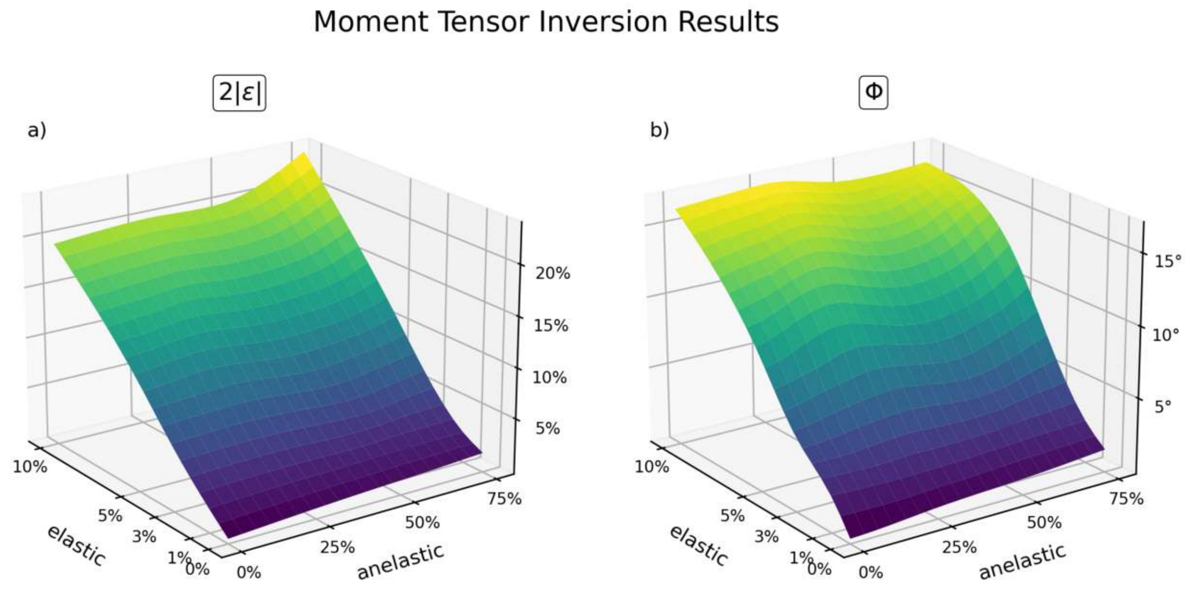


Figure 6 Inversion results for five inversions of each of the 25 earthquakes with five different perturbations of the elastic, and four different inversions of the anelastic structure. a) The resulting non-double-couple (NDC) components (2ϵ) depend primarily on the perturbation of the elastic Earth structure, with anelastic structure having little influence on the size of the NDC components. b) The angle required to rotate one moment tensor's set of principal axes into another (Φ) shows similar dependence on perturbations of the elastic and anelastic structure. The fault geometry is retrieved well, with rotation angles generally being smaller than $20°$.

nitude (Rösler et al., 2021). However, the NDC components vary significantly in polarity and in size between -17.1 and 5.6%.

Carrying out five inversions for each of the 25 earthquakes with five different perturbations of the elastic and four perturbations of the anelastic structure each yields 2500 inversions. The resulting NDC components depend primarily on the perturbation of the elastic structure of the Earth model (Fig. 6a). Similarly, the deviation in fault geometry (Fig. 6b) depends little on the perturbation of the anelastic structure, further illustrating the relative importance of elastic and anelastic Earth structure on seismic waveforms (Dahlen, 1982). The values for the rotation angle are generally small, indicating that the DC component is recovered relatively accurately by inversions that poorly represent the actual Earth structure, whereas the NDC components have large uncertainties and are often artifacts of the inversion.

The NDC components reported in the GCMT catalog average 23.5% for all earthquakes. However, earthquakes with M_w > 6.5 have, on average, smaller NDC components (17.2%), which are determined with greater precision (Rösler et al., 2021, 2023). The NDC components in our experiment resulting from inadequate representation of the Earth structure in the inversion are generally smaller than those of the earthquakes in the GCMT catalog when perturbing the elastic (Fig. 7a) and the anelastic Earth structure (Fig. 7b) independently. However, a perturbation of 10% in the elastic structure and 75% in the anelastic structure together reproduces the mean and standard deviation of the distribution of the NDC components in the GCMT catalog (Fig. 7c), making it possible to explain the NDC components in the GCMT catalog as resulting from not accounting for laterally varying Earth structure. Perturbing the elastic structure of the Earth model alone by 10% without perturbing the anelastic structure results in an average NDC component of 19.5%, only slightly smaller than the observed NDC components in the GCMT catalog. However, uncertainties in the anelastic structure are large (Karaoğlu and Romanowicz, 2018) and can reach 75% for seismic waves with periods of 100s as used in this experiment (Dalton and Ekström, 2006).

4 Discussion and Conclusions

Generating synthetic seismograms for the DC components of 25 arbitrarily selected earthquakes in the GCMT catalog using one Earth model and inverting them with another Earth model gives rise to MTs with NDC components because the inversion changes the mechanism to include a substantial NDC component. Perturbing both elastic and anelastic Earth structure yields a distribution of NDC components similar to that of NDC components reported for the earthquakes in the GCMT catalog. This process shows the sensitivity of MTs to the effects of variable Earth structure. This behavior is expected to be similar, but larger, for shorter periods than the > 100 s we present, which is presumably why smaller earthquakes have larger NDC components than larger earthquakes in global MT catalogs.

These results for global datasets are generally similar to those derived for regional data. Stierle et al. (2014b) found that unconstrained MT inversions allowing an isotropic component produce larger NDC compo-

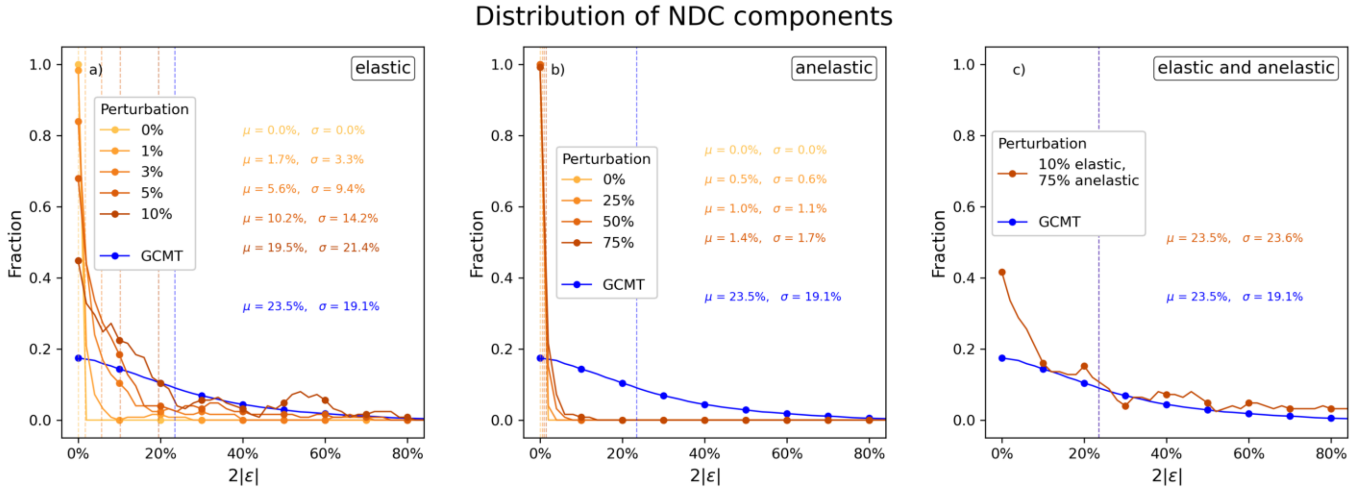


Figure 7 Distribution of non-double couple (NDC) components generated by inverting seismograms generated using a model with perturbed elastic (a) and anelastic (b) structure using unperturbed Preliminary Reference Earth Model (PREM). A combined perturbation of elastic and anelastic structure (c) generally reproduces the distribution of NDC components in the Global Centroid Moment Tensor (GCMT) catalog.

nents than constrained inversions. Stierle et al. (2014a) found that the aftershocks of the M_w 7.4 Izmit earthquake in 1999 had average NDC components of 19.6%. For the earthquakes of a swarm in 1997 in Czech Republic, Vavryčuk (2002) and Horálek et al. (2002) reported a mean deviation from a DC source of 17.3%. These studies found earthquakes with NDC components of up to 57.0% and 49.8%, respectively. The MT inversion in both studies used Green's functions generated for regional 1D Earth models, using the P- and S-wave amplitudes to stabilize the inversion, thus giving confidence that these NDC components represent real source processes.

NDC components of some earthquakes reflect complex source processes differing from slip on a planar fault or the combined effect of DC sources on multiple faults with different geometries. However, the fact that our numerical experiment replicates general features of the pervasive NDC components reported in moment tensor catalogs implies that these components are largely artifacts of the inversions not adequately accounting for the effects of laterally varying Earth structure.

This effect seems similar for the other MT catalogs we examined, which show comparable NDC components. The MTs in all global catalogs were derived using the one-dimensional Earth model PREM. However, the GCMT catalog corrects seismic waveforms for laterally varying Earth structure along the great-circle path of surface waves (Dziewonski et al., 1984). This approach changes the phase spectra, but not the amplitude spectra, and thus does not fully represent the expected effects of 3D structure.

NDC components are often attributed geologic meaning based on their size (Vavryčuk, 2002; Stierle et al., 2014a) without further investigation about their origin. Large NDC components and NDC components of large earthquakes are in fact more reliably determined in MT inversions based on 1D Earth models (Rösler et al., 2023) and are thus more likely to represent real source processes. However, based on the results of our numerical experiment, the threshold above which they can be considered real source processes based only on their size must be placed at 2σ from their global average of 23.5% at 61.7%, consistent with the results of Rösler et al. (2023). Earthquakes with real, but smaller NDC components exist, but require further knowledge about the geologic setting of the fault rupture occurred on, or knowledge from multiple MT inversions with different Earth models to confirm the significance of NDC components. MT inversions of global catalogs can be improved and artifactual NDC components reduced by using Green's functions generated for a 3D Earth model. However, Šílený and Vavryčuk (2002) found that isotropic and compensated linear-vector dipole (CLVD) components are overestimated for DC sources when inverting events with waveforms recorded in anisotropic structures but assuming isotropy. Therefore, the best estimates of NDC components require a laterally varying Earth model including anisotropy (Hjörleifsdóttir and Ekström, 2010; Sawade et al., 2022).

The results here, combined with the poor correlation between NDC components in different catalogs, suggest that the pervasive NDC components reported in moment tensor catalogs are largely artifacts of the inversions not adequately accounting for the effects of laterally varying Earth structure. More realistic estimates of NDC components will thus require inversion methods that better model the effects of lateral variability which increase computational cost and are less applicable for routinely determined MTs.

Acknowledgements

We thank Emile Okal, Craig Bina, Reece Elling, James Neely, and Molly Gallahue for helpful discussions that improved the manuscript, and for the suggestions made by anonymous reviewers and the handling editor, Stephen Hicks. We thank Brian Shiro and David Wilson for comments and handling the USGS approval of this manuscript, and Dara Goldberg and Janet Carter for providing internal reviews. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Data and code availability

The Global CMT Project catalog used in this study was downloaded from globalcmt.org. We used the Mineos software package (Masters et al.), available at geodynamics.org, and station locations of the Global Seismographic Network to generate synthetic seismograms. These data are freely available from the Earthscope Consortium. For the data analysis, we used ObsPy (Beyreuther et al., 2010), and BayesISOLA for the moment tensor inversion, available at github.com/vackar/BayesISOLA.

The EarthScope Consortium is funded through the Seismological Facility for the Advancement of Geoscience (SAGE) Award of the National Science Foundation under Cooperative Agreement EAR-1724509.

The Global Seismographic Network (GSN) is a cooperative scientific facility operated jointly by the Earthscope Consortium, the U.S. Geological Survey (USGS), and the National Science Foundation (NSF), under Cooperative Agreement EAR-1261681.

Competing interests

The authors declare no competing interests.

References

Aki, K. and Richards, P. G. Quantitative Seismology. *University Science Books*, 2002.

Ammon, C. J., Lay, T., Velasco, A. A., and Vidale, J. E. Routine estimation of earthquake source complexity: The 18 October 1992 Colombian earthquake. *Bulletin of the Seismological Society of America*, 4(84):1266–1271, 1994. doi: 10.1785/BSSA0840041266.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., , and Wassermann, J. ObsPy: A Python toolbox for seismology. *Seismological Research Letters*, 81(3):530–533, 2010. doi: 10.1785/gssrl.81.3.530.

Cesca, S., Buforn, E., and Dahm, T. Amplitude spectra moment tensor inversion of shallow earthquakes in Spain. *Geophysical Journal International*, 166(2):839–854, 2006. doi: 10.1111/j.1365-246X.2006.03073.x.

Cohee, B. P. and Beroza, G. C. Slip distribution of the 1992 Landers earthquake and its implications for earthquake source mechanics. *Bulletin of the Seismological Society of America*, 84(3): 692–712, 1994. doi: 10.1785/BSSA0840030692.

Covellone, B. M. and Savage, B. A quantitative comparison between 1D and 3D source inversion methodologies: Application to the Middle East. *Bulletin of the Seismological Society of America*, 102(5):2189–2199, 2012. doi: 10.1785/0120110278.

Dahlen, F. and Tromp, J. Theoretical global seismology. *Princeton University Press*, 1998.

Dahlen, F. A. The effect of data windows on the estimation of free oscillation parameters. *Geophysical Journal International*, 2(69):537–549, 1982. doi: 10.1111/j.1365-246X.1982.tb04964.x.

Dalton, C. A. and Ekström, G. Global models of surface wave attenuation. *Journal of Geophysical Research: Solid Earth*, B5(111), 2006. doi: 10.1029/2005JB003997.

Domingues, A., Custodio, S., and Cesca, S. Waveform inversion of small-to-moderate earthquakes located offshore southwest Iberia. *Geophysical Journal International*, 192(1):248–259, 2013. doi: 10.1093/gji/ggs010.

Donner, S., Mustać, M., Hejrani, B., Tkalčić, H., and Igel, H. Seismic moment tensors from synthetic rotational and translational ground motion: Green's functions in 1-D versus 3-D. *Geophysical Journal International*, 223(1):161–179, 2020. doi: 10.1093/gji/ggaa305.

Duputel, Z., Rivera, L., Kanamori, H., and Hayes, G. W phase source inversion for moderate to large earthquakes (1990-2010). *Geophysical Journal International*, 2(189):1125–1147, 2018. doi: 10.1111/j.1365-246X.2012.05419.x.

Dziewonski, A. M. and Anderson, D. L. Preliminary reference Earth model. *Physics of the Earth and Planetary Interiors*, 25(4): 297–356, 1981. doi: 10.1016/0031-9201(81)90046-7.

Dziewonski, A. M., Chou, T. A., and Woodhouse, J. H. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *Journal of Geophysical Research: Solid Earth*, B4(86):2825–2852, 1981. doi: 10.1029/JB086iB04p02825.

Dziewonski, A. M., Franzen, J. E., and Woodhouse, J. H. Centroid-moment tensor solutions for January-March, 1984. *Physics of the Earth and Planetary Interiors*, 4(34):209–219, 1984. doi: 10.1016/0031-9201(84)90062-1.

Ekström, G., Nettles, M., and Dziewonski, A. The global CMT project 2004-2010: Centroid-moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors*, 200-201:1–9, 2012. doi: 10.1016/j.pepi.2012.04.002.

Ford, S. R., Dreger, D. S., and Walter, W. R. Network sensitivity solutions for regional moment-tensor inversions. *Bulletin of the Seismological Society of America*, 100(5A):1962–1970, 2010. doi: 10.1785/0120090140.

Frohlich, C. Triangle diagrams: ternary graphs to display similarity and diversity of earthquake focal mechanisms. *Physics of the Earth and Planetary Interiors*, 75(1-3):193–198, 1992. doi: 10.1016/0031-9201(92)90130-N.

Frohlich, C. Earthquakes with non-double-couple mechanisms. *Science*, 5160(264):804–809, 1994. doi: 10.1126/science.264.5160.804.

Gilbert, F. Excitation of the normal modes of the Earth by earthquake sources. *Geophysical Journal International*, 2(22): 223–226, 1971. doi: 10.1111/j.1365-246X.1971.tb03593.x.

Gudmundsson, M. T., Jónsdóttir, K., A., H., Holohan, E. P., Halldórsson, S. A., Ófeigsson, B. G., Cesca, S., Vogfjörd, K. S., Sigmundsson, F., Högnadóttir, T., Einarsson, P., Sigmarsson, O., Jarosch, A. H., Jónasson, K., Magnússon, E., Hreinsdóttir, S., Bagnardi, M., Parks, M. M., Hjörleifsdóttir, V., Pálsson, F., W. T. R., Schöpfer, M. P. J., Heimann, S., Reynolds, H. I., Dumont, S., Bali, E., Gudfinnsson, G. H., Dahm, T., Roberts, M. J., Hensch, M., Belart, J. M. C., Spaans, K., Jakobsson, S., Gudmundsson, G. B., Fridriksdóttir, H. M., Drouin, V., Dürig, T., Adalgeirsdóttir, G., Riishuus, M. S., Pedersen, G. B. M., Van Boeckel, T., Oddsson, B., Pfeffer, M. A., Barsotti, S., Bergsson, B., Donovan, A., Burton, M. R., and Aiuppa, A. Gradual caldera collapse at Bárdarbunga volcano, Iceland, regulated by lateral magma outflow. *Science*, 353 (6296), 2016. doi: 10.1126/science.aaf8988.

Hallo, M. and Gallovič, F. Fast and cheap approximation of Green function uncertainty for waveform-based earthquake source inversions. *Geophysical Journal International*, 207(2):1012–1029, 2016. doi: 10.1093/gji/ggw320.

Hamling, I. J., Hreinsdóttir, S., Clark, K., Elliott, J., Liang, C., Fielding, E., Litchfield, N., Villamor, P., Wallace, L., Wright, T. J., d'Anastasio, E., Bannister, S., Burbridge, D., Denys, P., Gentle, P., Howarth, J., Mueller, C., Palmer, N., Pearson, C., Power, W., Barnes, P., Barrell, D. J. A., van Dissen, R., Langridge, R., Little, T., Nicol, A., Pettinga, J., Rowland, J., and Stirling, M. Complex multifault rupture during the M_w 7.8 Kaikōura earthquake, New Zealand. *Science*, 356(6334), 2017. doi: 10.1126/science.aam7194.

Hayes, G. P., Rivera, L., and Kanamori, H. Source inversion of the W-Phase: real-time implementation and extension to low magnitudes. *Seismological Research Letters*, 80(5):817–822, 2009. doi: 10.1785/gssrl.80.5.817.

Hayes, G. P., Briggs, R. W., Sladen, A., Fielding, E. J., Prentice, C., Hudnut, K., Mann, P., Taylor, F. W., Crone, A. J., Gold, R., Ito, T., and Simons, M. Complex rupture during the 12 January 2010 Haiti earthquake. *Nature Geoscience*, 3(11):800–805, 2010. doi: 10.1038/ngeo977.

Hejrani, B., Tkalčić, H., and Fichtner, A. Centroid moment tensor catalogue using a 3-D continental scale Earth model: Application to earthquakes in Papua New Guinea and the Solomon Islands. *Journal of Geophysical Research: Solid Earth*, 122(7): 5517–5543, 2017. doi: 10.1002/2017JB014230.

Henry, C., Woodhouse, J. H., and S., D. Stability of earthquake moment tensor inversions: effect of the double-couple constraint. *Tectonophysics*, 356(1-3):115–124, 2002. doi: 10.1016/S0040-1951(02)00379-7.

Hingee, M., Tkalčić, H., Fichtner, A., and Sambridge, M. Seismic moment tensor inversion using a 3-D structural model: applications for the Australian region. *Geophysical Journal International*, 184(2):949–964, 2011. doi: 10.1111/j.1365-246X.2010.04897.x.

Hjörleifsdóttir, V. and Ekström, G. Effects of three-dimensional Earth structure on CMT earthquake parameters. *Physics of the Earth and Planetary Interiors*, 179(3-4):178–190, 2010. doi: 10.1016/j.pepi.2009.11.003.

Horálek, J., Šílený, J., and Fischer, T. Moment tensors of the January 1997 earthquake swarm in NW Bohemia (Czech Republic): double-couple vs. non-double-couple events. *Tectonophysics*, 1-3(356):65–85, 2002. doi: 10.1016/S0040-1951(02)00377-3.

Jechumtálová, Z. and Bulant, P. Effects of 1-D versus 3-D velocity models on moment tensor inversion in the Dobrá Voda area in the Little Carpathians region, Slovakia. *Journal of Seismology*, 18:511–531, 2014. doi: 10.1007/s10950-014-9423-6.

Jechumtálová, Z. and Šílený, J. Point-source parameters from noisy waveforms: Error estimate by Monte-Carlo simulation. *Pure and Applied Geophysics*, 158(9):1639–1654, 2001. doi: 10.1007/PL00001237.

Julian, B. R. and Sipkin, S. A. Earthquake processes in the Long Valley caldera area, California. *Journal of Geophysical Research: Solid Earth*, 90(B13):11155–11169, 1985. doi: 10.1029/JB090iB13p11155.

Kagan, Y. Y. 3-D rotation of double-couple earthquake sources. *Geophysical Journal International*, 106(3):709–716, 1991. doi: 10.1111/j.1365-246X.1991.tb06343.x.

Kanamori, H. and Given, J. W. Analysis of long-period seismic waves excited by the May 18, 1980, eruption of Mount St. Helens - A terrestrial monopole? *Journal of Geophysical Research: Solid Earth*, 87(B7):5422–5432, 1982. doi: 10.1029/JB087iB07p05422.

Kanamori, H. and Rivera, L. W phase. *Geophysical Research Letters*, 16(20):1691–1694, 1993. doi: 10.1029/93GL01883.

Kanamori, H., R. L. Source inversion of W-phase: Speeding up seismic tsunami warning. *Geophysical Journal International*, 1(175): 222–238, 2008. doi: 10.1111/j.1365-246X.2008.03887.x.

Karaoğlu, H. and Romanowicz, B. Inferring global upper-mantle shear attenuation structure by waveform tomography using the spectral element method. *Geophysical Journal International*, 3 (213):1536–1558, 2018. doi: 10.1093/gji/ggy030.

Kawakatsu, H. Enigma of earthquakes at ridge-transform-fault plate boundaries distribution of non-double-couple parameter of Harvard CMT Solutions. *Geophysical Research Letters*, 18(6): 1103–1106, 1991. doi: 10.1029/91GL01238.

Knopoff, L. and Randall, M. J. The compensated linear-vector dipole: A possible mechanism for deep earthquakes. *Journal of Geophysical Research*, 26(75):4957–4963, 1970. doi: 10.1029/JB075i026p04957.

Liu, Q., Polet, J., Komatitsch, D., and Tromp, J. Spectral-element moment tensor inversions for earthquakes in southern California. *Bulletin of the Seismological Society of America*, 94(5): 1748–1761, 2004. doi: 10.1785/012004038.

Masters, G., Woodhouse, J., and Freeman, G. *Computational Infrastructure for Geodynamics, url: https://geodynamics.org/cig*.

McNamara, D. E., Benz, H. M., Herrmann, R. B., Bergman, E. A., Earle, P., Meltzer, A., Withers, M., and Chapman, M. The Mw 5.8 Mineral, Virginia, Earthquake of August 2011 and Aftershock Sequence: Constraints on Earthquake Source Parameters and Fault Geometry. *Bulletin of the Seismological Society of America*, 1(104):40–54, 2013.

Miller, A. D., Foulger, G. R., and Julian, B. R. Non-double-couple earthquakes 2. Observations. *Reviews of Geophysics*, 36(4): 551–568, 1998. doi: 10.1029/98RG00717.

Nettles, M. and Ekström, G. Faulting mechanism of anomalous earthquakes near Bárdarbunga Volcano, Iceland. *Journal of Geophysical Research: Solid Earth*, 103(B8):17973–17983, 1998. doi: 10.1029/98JB01392.

Pang, G., Koper, K. D., Mesimeri, M., Pankow, K. L., Baker, B., Farrell, J., Holt, J., Hale, J. M., Roberson, P., Burlacu, R., Pechmann, J. C., Whidden, K., Holt, M. M., Allamm, A., and DuRoss, C. Seismic analysis of the 2020 Magna, Utah, earthquake sequence: Evidence for a listric Wasatch fault. *Geophysical Research Letters*, 47(18), 2020. doi: 10.1029/2020GL089798.

Phạm, T. S. and Tkalčić, H. Toward Improving Point-Source Moment-Tensor Inference by Incorporating 1D Earth Model's Uncertainty: Implications for the Long Valley Caldera Earthquakes. *Journal of Geophysical Research: Solid Earth*, 126(11), 2021. doi: 10.1029/2021JB022477.

Quigley, M. C., Mohammadi, H., and Duffy, B. G. Multi-fault earthquakes with kinematic and geometric rupture complexity: how common? *INQUA Focus Group Earthquake Geology and Seismic Hazards*, 2017.

Ringler, A. T., Anthony, R. E., Aster, R. C., Ammon, C. J., Arrowsmith, S., Benz, H., Ebeling, C., Frassetto, A., Kim, W.-Y., Koelemeijer, P., Lau, H. C. P., Lekić, V., Montagner, J. P., Richards, P. G., Schaff, D. P., Vallee, M., and Yeck, W. Achievements and prospects of global broadband seismographic networks after 30 years of continuous geophysical observations. *Reviews of Geophysics*, 3(60), 2022. doi: 10.1029/2021RG000749.

Rodríguez-Cardozo, F., Hjörleifsdóttir, V., K., J., Iglesias, A., Franco, S. I., Geirsson, H., Trujillo-Castrillón, N., and M., H. The 2014-2015 complex collapse of the Bárðarbunga caldera, Iceland, revealed by seismic moment tensors. *Journal of Volcanology and Geothermal Research*, 416, 2021. doi: 10.1016/j.jvolgeores.2021.107275.

Rösler, B. and Stein, S. Consistency of Non-Double-Couple Components of Seismic Moment Tensors with Earthquake Magnitude and Mechanism. *Seismological Research Letters*, 93(3): 1510–1523, 2022. doi: 10.1785/0220210188.

Rösler, B., Stein, S., and Spencer, B. D. Uncertainties in Seismic

Moment Tensors Inferred from Differences Between Global Catalogs. *Seismological Research Letters*, 92(6):3698–3711, 2021. doi: 10.1785/0220210066.

Rösler, B., Stein, S., and Spencer, B. D. When are Non-Double-Couple Components of Seismic Moment Tensors Reliable? *Seismica*, 2(1), 2023.

Ross, A., Foulger, G. R., and Julian, B. R. Non-double-couple earthquake mechanisms at the Geysers geothermal area, California. *Geophysical Research Letters*, 23(8):877–880, 1996. doi: 10.1029/96GL00590.

Rößler, D., Krüger, F., and Rümpker, G. Retrieval of moment tensors due to dislocation point sources in anisotropic media using standard techniques. *Geophysical Journal International*, 169(1): 136–148, 2007. doi: 10.1111/j.1365-246X.2006.03243.x.

Ruhl, C. J., Morton, E. A., Bormann, J. M., Hatch-Ibarra, R., Ichinose, G., and Smith, K. D. Complex Fault Geometry of the 2020 M_{ww} 6.5 Monte Cristo Range, Nevada, Earthquake Sequence. *Seismological Research Letters*, 92(3):1876–1890, 2021. doi: 10.1785/0220200345.

Saloor, N. and Okal, E. A. Extension of the energy-to-moment parameter Θ to intermediate and deep earthquakes. *Physics of the Earth and Planetary Interiors*, 274:37–48, 2018. doi: 10.1016/j.pepi.2017.10.006.

Sandanbata, O., Kanamori, H., Rivera, L., Zhan, Z., Watada, S., and Satake, K. Moment tensors of ring-faulting at active volcanoes: Insights into vertical-CLVD earthquakes at the Sierra Negra caldera, Galápagos Islands. *Journal of Geophysical Research: Solid Earth*, 126(6), 2021. doi: 10.1029/2021JB021693.

Sawade, L., Beller, S., L. W., and Tromp, J. Global centroid moment tensor solutions in a heterogeneous earth: the CMT3D catalogue. *Geophysical Journal International*, 231(3):1727–1738, 2022. doi: 10.1093/gji/ggac280.

Scognamiglio, L., Tinti, E., Casarotti, E., Pucci, S., Villani, F., Cocco, M., Magnoni, F., Michelini, A., and Dreger, D. Complex fault geometry and rupture dynamics of the M_w 6.5, 30 October 2016, Central Italy earthquake. *Journal of Geophysical Research: Solid Earth*, 123(4):2943–2964, 2018. doi: 10.1002/2018JB015603.

Shuler, A., Ekström, G., and Nettles, M. Physical mechanisms for vertical-CLVD earthquakes at active volcanoes. *Journal of Geophysical Research: Solid Earth*, 118(4):1569–1586, 2013a. doi: 10.1002/jgrb.50131.

Shuler, A., Nettles, M., and Ekström, G. Global observation of vertical-CLVD earthquakes at active volcanoes. *Journal of Geophysical Research: Solid Earth*, 118(1):138–164, 2013b. doi: 10.1029/2012JB009721.

Šílený, J. Regional moment tensor uncertainty due to mismodeling of the crust. *Tectonophysics*, 383(3-4):133–147, 2004. doi: 10.1016/j.tecto.2003.12.007.

Šílený, J., Campus, P., and Panza, G. F. Seismic moment tensor resolution by waveform inversion of a few local noisy records - I. Synthetic tests. *Geophysical Journal International*, 126(3): 605–619, 1996. doi: 10.1111/j.1365-246X.1996.tb04693.x.

Sipkin, S. A. Interpretation of non-double-couple earthquake mechanisms derived from moment tensor inversion. *Journal of Geophysical Research: Solid Earth*, B1(91):531–547, 1986. doi: 10.1029/JB091iB01p00531.

Stierle, E., Bohnhoff, M., and Vavryčuk, V. Resolution of non-double-couple components in the seismic moment tensor using regional networks-II: application to aftershocks of the 1999 Mw 7.4 Izmit earthquake. *Geophysical Journal International*, 3 (169):1878–1888, 2014a. doi: 10.1093/gji/ggt503.

Stierle, E., Vavryčuk, V., Šílený, J., and Bohnhoff, M. Resolution of non-double-couple components in the seismic moment tensor using regional networks-I: a synthetic case study. *Geophysical Journal International*, 3(169):1869–1877, 2014b. doi: 10.1093/gji/ggt502.

Tréhu, A. M., Nábělek, J. L., and Solomon, S. C. Source characterization of two Reykjanes Ridge earthquakes: Surface waves and moment tensors; P waveforms and nonorthogonal nodal planes. *Journal of Geophysical Research: Solid Earth*, B3(86): 1701–1724, 1981. doi: 10.1029/JB086iB03p01701.

Vackář, Burjánek, J., Gallovič, F., Zahradník, J., and Clinton, J. Bayesian ISOLA: new tool for automated centroid moment tensor inversion. *Geophysical Journal International*, 2(210): 693–705, 2017. doi: 10.1093/gji/ggx158.

Vasyura-Bathke, H., Dettmer, J., Dutta, R., Mai, P. M., and Jonsson, S. Accounting for theory errors with empirical Bayesian noise models in nonlinear centroid moment tensor estimation. *Geophysical Journal International*, 225(2):1412–1431, 2021. doi: 10.1093/gji/ggab034.

Vavryčuk, V. Non-double-couple earthquakes of 1997 January in West Bohemia, Czech Republic: evidence of tensile faulting. *Geophysical Journal International*, 2(149):364–373, 2002. doi: 10.1046/j.1365-246X.2002.01654.x.

Vera Rodriguez, I., Gu, Y. J., and Sacchi, M. D. Resolution of seismic-moment tensor inversions from a single array of receivers. *Bulletin of the Seismological Society of America*, 101(6):2634–2642, 2011. doi: 10.1785/0120110016.

Šílený, J. and Vavryčuk, V. Can unbiased source be retrieved from anisotropic waveforms by using an isotropic model of the medium? *Tectonophysics*, 1-3(356):125–138, 2002. doi: 10.1016/S0040-1951(02)00380-3.

Wald, D. J. and Heaton, T. H. Spatial and temporal distribution of slip for the 1992 Landers, California, earthquake. *Bulletin of the Seismological Society of America*, 84(3):668–691, 1994. doi: 10.1785/BSSA0840030668.

Wang, X. and Zhan, Z. Moving from 1-D to 3-D velocity model: automated waveform-based earthquake moment tensor inversion in the Los Angeles region. *Geophysical Journal International*, 220 (1):218–234, 2020. doi: 10.1093/gji/ggz435.

Yang, J., Zhu, H., Lay, T., Niu, Y., Ye, L., Lu, Z., Luo, B., Kanamori, H., Huang, J., and Li, Z. Multi-fault opposing-dip strike-slip and normal-fault rupture during the 2020 M_w 6.5 Stanley, Idaho earthquake. *Geophysical Research Letters*, 48(10), 2021. doi: 10.1029/2021GL092510.

Zhu, L. and Zhou, X. Seismic moment tensor inversion using 3D velocity model and its application to the 2013 Lushan earthquake sequence. *Physics and Chemistry of the Earth, Parts A/B/C*, 95: 10–18, 2016. doi: 10.1016/j.pce.2016.01.002.

doi:10.26443/seismica.v3i1.1145

SEISMICA

Spatiotemporal characteristics and earthquake statistics of the 2020 and 2022 adjacent earthquake sequences in North Aegean Sea (Greece)

P. Bonatis, * [1], **V. Karakostas**, [1], **C. Kourouklas**, [1], **A. Kostoglou**, [1,2], **E. Papadimitriou** [1]

[1]Geophysics Department, Faculty of Geology, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece, [2]Institute of Geophysics, Polish Academy of Sciences, Department of Seismology, Ks. Janusza 64 Street, 01-452 Warsaw, Poland

Author contributions: *Conceptualization*: P. Bonatis, V. Karakostas, A. Kostoglou. *Data Curation*: P. Bonatis, V. Karakostas, C. Kourouklas. *Formal Analysis*: P. Bonatis, C. Kourouklas, A. Kostoglou. *Investigation*: All authors. *Methodology*: P. Bonatis, V. Karakostas, C. Kourouklas. *Resources*: V. Karakostas, E. Papadimitriou. *Software*: P. Bonatis, V. Karakostas, C. Kourouklas, A. Kostoglou. *Supervision*: V. Karakostas, E. Papadimitriou. *Validation*: V. Karakostas, E. Papadimitriou. *Visualization*: P. Bonatis, C. Kourouklas. *Writing – original draft*: P. Bonatis, C. Kourouklas, A. Kostoglou, E. Papadimitriou. *Writing – review & editing*: All authors.

Abstract The two moderate earthquakes that occurred close and to the north of the North Aegean Trough (NAT) on 26 September 2020 (M_w5.3) and 16 January 2022 (M_w5.4), both followed by aftershock activity, are examined. Seismic activity along the NAT and its parallel branches is continuous and remarkable, with numerous strong instrumental (M≥6.0) earthquakes. Yet, the frequency of moderate (5.0≤M<6.0) earthquakes outside these major fault branches is rather rare and therefore their investigation provides the optimal means to decipher the seismotectonic properties of the broader area. The temporal and spatial proximity of the two seismic excitations from late September of 2020 through early 2022, intrigues for exhaustive investigation of seismic activity with the employment of earthquake relocation techniques, moment tensor solutions and statistical analysis. Our research revealed that this seismic activity purely falls inside the Mainshock – Aftershock type, with fast aftershock decay rates and moderate productivity. According to our findings, the two seismic sequences, despite their close proximity, exhibit distinctive features as a result of the intricate stress field generated at the western termination of the NAF system in an extensional domain.

Production Editor:
Gareth Funning
Handling Editor:
Yen Joe Tan
Copy & Layout Editor:
Jack B. Muir

Signed reviewer(s):
Kostas Leptokaropoulos

Received:
2023-11-03
Accepted:
2024-03-19
Published:
2024-05-08

Περίληψη Αντικείμενο της μελέτης μας είναι οι δύο σεισμοί ενδιαμέσου μεγέθους που έγιναν πλησίον και βόρεια της τάφρου του Β. Αιγαίου στις 26 Σεπτεμβρίου 2020 (M_w5.3) και στις 16 Ιανουαρίου 2022 (M_w5.4) μαζί με τις μετασεισμικές τους ακολουθίες. Η σεισμική δραστηριότητα κατά μήκος της τάφρου του Β. Αιγαίου και των παράλληλων κλάδων της δεν είναι σπάνια, με πολλούς ισχυρούς σεισμούς (M≥6.0) να έχουν καταγραφεί κατά την ενόργανη περίοδο της σεισμικότητας. Από την άλλη πλευρά, ενδιαμέσου μεγέθους (5.0≤M<6.0) σεισμοί εκτός αυτών των κύριων κλάδων συμβαίνουν πολύ σπάνια, ως εκ τούτου η μελέτη τους αποτελεί μία πρώτης τάξης ευκαιρία για την αναλυτική ερμηνεία των σεισμοτεκτονικών ιδιοτήτων της ευρύτερης περιοχής. Η χωρική και χρονική εγγύτητα των δύο σεισμικών εξάρσεων που τέθηκε σε εξέλιξη από τα τέλη Σεπτεμβρίου του 2022 έως τις αρχές του 2022 ευνοεί την ενδελεχή διερεύνηση της σεισμικής δραστηριότητας με τη χρήση τεχνικών για τον επαναπροσδιορισμό των εστιακών συντεταγμένων, τον καθορισμό των μηχανισμών γένεσης τους και στατιστική ανάλυση της σεισμικότητας. Τα αποτελέσματα της στατιστικής ανάλυσης που πραγματοποιήθηκε υποδεικνύει ότι οι σεισμικές αυτές εξάρσεις ακολουθούν τυπικό μοτίβο ακολουθίας Κύριος σεισμός–Μετασεισμοί, με υψηλή απόσβεση των ρυθμών σεισμικότητας και μέτριου βαθμού παραγωγικότητα. Σύμφωνα με τα ευρήματά μας οι δύο σεισμικές ακολουθίες, παρά τη χωροχρονική τους εγγύτητα παρουσιάζουν διακριτά χαρακτηριστικά ως αποτέλεσμα του περίπλοκου τοπικού καθεστώτος τάσεων που οφείλεται στον τερματισμό της ζώνης μετασχηματισμού της Β. Ανατολίας σε ένα εφελκυστικό καθεστώς.

Non-technical summary On 26 September 2020 and 16 January 2022 two moderate earthquakes (M_w5.3 and M_w5.4, respectively) occurred at the North Aegean Sea, southern of Chalkidiki peninsula. Their close proximity in space and time and the rare manifestation of such moderate events in the area promotes their analysis in order to better understand the faults and the state of stress in the broader area. We used the available seismological stations in the area to enhance the quality of the earthquake catalog and thoroughly investigated the properties of the two seismic sequences. We found that this activity is not directly related to the prevailing seismotectonic feature of the area, namely the North Aegean Trough, but it is derived from secondary features, typically found in the vicinity of such large complex systems.

*Corresponding author: mponatis@geo.auth.gr

1 Introduction

The rollback of the oceanic lithospheric plate of eastern Mediterranean that subducts beneath the continental crust of the Aegean microplate (Pichon and Angelier, 1979; Papazachos and Comninakis, 1971) is the driving mechanism for the fast extensional deformation in an almost N–S direction of the back arc region (Konstantinou et al., 2017; Kapetanidis and Kassaras, 2019). The westward prolongation of the North Anatolian Fault (NAF) into the Aegean forms the North Aegean Trough (NAT) that constitutes the active boundary between the Aegean microplate and the Eurasian lithospheric plate (inset of Figure 1). McKenzie (1972) showed that the northward motion of the Arabian plate pushes the smaller Anatolian plate that is moving westerly relative to the Eurasian plate along the NAF, with an average velocity of about ~24 mm/yr. An additional N–S deformation of ~11 mm/yr in the Aegean further enhances this motion, resulting in a total SW motion of ~41 mm/yr of the south Aegean relative to Eurasia. More recent studies (e.g. England et al., 2016; Bitharis et al., 2023) comply with this overall pattern. A large part of the deformation occurs seismically, as the Aegean shows a total seismic slip rate of the order of 20 mm/yr relative to Eurasia (Papazachos and Kiratzi, 1996). Seismicity is intense and continuous in the back arc region forming specific seismic zones striking almost E–W in the normal faulting environment and ENE–WSW to NE–SW for the dextral strike-slip and oblique f).aulting seismic zones (Papazachos et al., 1998). Conjugate faulting sinistral strike-slip faults are also present (Karakostas et al., 2003) accommodating less frequent M≥6.0 main shocks as well as moderate (M≥5.0) earthquakes.

Several strong (M≥6.0) historical and instrumental earthquakes have struck the NAT (Figure 1) since 360 BC (Papazachos and Papazachou, 2003). Their temporal distribution evidences catalog incompleteness at least until 1300 AD, and afterward missing events are the ones of M<6.5 (Kourouklas et al., 2018). Most recently, the 2014 M6.9 main shock ruptured the middle part of the NAT (Kiratzi et al., 2016), taking place in a previously identified seismic gap (Karakostas et al., 1986). The aftershock activity was astonishingly weak with a maximum magnitude aftershock of M4.9. However, the western termination of NAT against the Greek mainland appears to be relatively quiescent (Kourouklas et al., 2018, 2022). The strike-slip faulting of the NAT is confirmed by the fault plane solutions of the stronger earthquakes in the most recent part of the instrumental era, when seismological networks were capable to provide adequate data, such as in the case of the 1983 seismic sequence (Rocca et al., 1985). The results from waveform modeling performed for the 1968 (M7.5; Pacheco and Sykes (1992)), 1981 (M6.8) and 1983 (M6.6) earthquakes established the dextral strike-slip nature of the major faults in North Aegean with the axis of maximum extension, T, striking N–S and being nearly horizontal (Kiratzi et al., 1991).

Because it is difficult to determine the causative faults of the moderate (M≥5.0) earthquakes, our understanding of the seismogenic setting of the secondary faults in our study area is rather limited. An M5.8 main shock occurred in 2013 on one of the ENE–WSW trending parallel dextral strike-slip fault branches in Northern Aegean, in the continuation of the 1968 large (M=7.5) rupture (Drakopoulos and Ekonomides, 1972). Its rich aftershock sequence was scrutinized and the off–fault seismicity was perfectly explained with Coulomb stress changes when the parameters of friction and Skempton's coefficients attained values of μ>0.5 and B=0.0, respectively, implying high fault friction (Karakostas et al., 2014) Thus, clarifying the properties of the causative faults of moderate earthquakes is crucial to reveal the seismogenic structures and their relation to the major fault zones and evaluate the subsequent seismic hazard.

Microseismicity studies are not common in the study area, given that active faults are off-shore, and near-fault instruments are not in operation. The western termination of NAT was investigated by Hatzfeld et al. (1999) who confirmed the strike–slip motion which has probably been active since the Pliocene. This strike–slip motion is transferred into normal faulting (with the direction of principal extension being constant) in continental Greece. Accurate relocation of seismicity during 2011–2016 along the NAT yielded a pick of the focal depth distribution at 8 km diminishing down to 20 km (Konstantinou, 2017). The relocated seismicity defines distinctive clusters, one of which is located inside our study area, but the spatial distribution is quite disperse and thus not adequate to identify a specific activated structure.

On 26 September 2020 (at 22:50:24 UTC), an M_w5.3 earthquake occurred (right magenta circle in Figure 1), to the north of the NE–SW trending western portion of NAT and very close to its inferred trace, in a way that it would be considered as associated with failure on a fault patch along the NAT interface. It occurred in the middle of the night and was widely felt, alarming plenty of citizens even at distances greater than 200 km (EMSC felt reports). The GCMT (Global Centroid Moment Tensor; https://www.globalcmt.org/) solution suggests strike-slip faulting with one of the nodal planes striking NE–SW that complies with the sense of motion along the NAT (focal mechanisms configured with magenta compressional quadrants in Figure 1). The appreciable aftershock activity, however, persistently aligned in a NW–SE direction in agreement with the strike of the second nodal plane. These first observations provoked our interest in investigating the characteristics of the activated structure. Then, fifteen months later, a second moderate M_w5.4 earthquake occurred on 16 January 2022 (at 11:48:06 UTC) very close to the epicenter of the first 2020 shock (left magenta circle in Figure 1). The 2022 aftershock activity exhibited the same pattern as in 2020, and the second main shock's GCMT solution implies oblique faulting with the one nodal plane also striking NW–SE. In this study, we relocate the aftershocks and determine the fault plane solutions of the strongest among them. We further study the spatiotemporal evolution of the activity and perform statistical analysis through the application of the Epidemic Type Aftershock Sequence (ETAS) model.

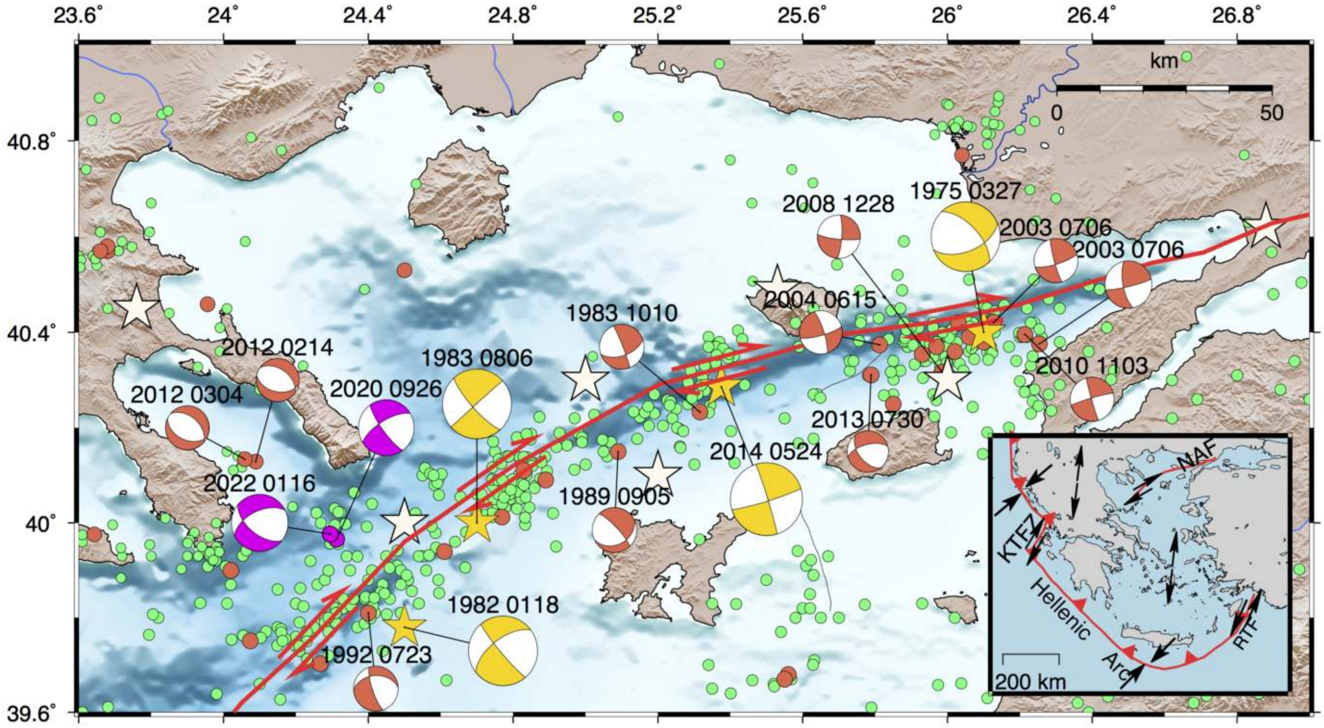


Figure 1 Epicentral distribution of M≥4.1 crustal (0≤h≤20 km) earthquakes occurred in northern Aegean Sea from 1975 to 2022 (after Leptokaropoulos et al., 2012), as compiled from the regional parametric earthquake catalog of the Seismological Station of Geophysics Department of the Aristotle University of Thessaloniki (http://geophysics.geo.auth.gr/ss/catalogs_en.html; Aristotle University of Thessaloniki, 1981). The small light green and moderate orange circles depict the epicenters of 4.1≤M<5.0 and 5.0≤M<6.0 earthquakes, respectively, and magenta circles the epicenters of for the earthquakes investigated in this study. White stars display the epicenters of the historical M≥6.5 earthquakes between 1845–1900. Yellow stars depict epicenters of the M≥6.0 earthquakes occurred during the instrumental era. The available fault plane solutions of the M≥5.0 earthquakes as taken from Global Centroid Moment Tensor (GCMT; https://www.globalcmt.org/) are plotted as equal area lower hemisphere projections with the compressional quadrants colored according to their magnitude as before. Solid thick red line represents the main branch of the North Aegean Trough Fault Zone, whereas the parallel red arrows imply the right-lateral strike-slip motion. Inset map shows the main seismotectonic features of the Aegean region (solid red lines; KTFZ: Kefalonia Transform Fault Zone; RTF: Rodos Transform Fault; NAF: North Anatolian Fault).

2 Seismological Data and Methods

2.1 Earthquake Relocation

An initial earthquake catalog was compiled by retrieving the recordings of the Hellenic Unified Seismological Network (HUSN) after routine analysis accomplished in the central Seismological Station of the Geophysics Department of the Aristotle University of Thessaloniki (http://geophysics.geo.auth.gr/ss/) for earthquakes with M≥2.5. Additional phase picking and initial location were performed by the authors of the present study to achieve the inclusion of lower magnitude events (even lower than M=1.0) that were detected by stations quite close to the activated area. In total, 841 earthquakes were initially located with magnitudes 0.6<M<5.4. For the earthquake relocation procedure, we used data from 14 seismological stations at epicentral distances up to approximately 150 km (Figures S1 & S2). The median azimuthal gap for all initial locations equals to 108° indicating a sufficient station azimuthal coverage, considering that the activated structures are offshore.

The first step towards improved focal coordinate estimation is relocation with the HYPOINVERSE code (Klein, 2002). This step requires the manually picked P– and S– phases, a local velocity model, the ratio of compressional to shear wave velocity (V_p/V_s) and the corresponding station time delays. At first, we estimated the V_p/V_s through the Wadati method (Wadati, 1933) by taking the minimum number of phase arrival pairs (P– and S–) equal to 8 for each event, and the process resulted in a ratio equal to 1.74 ± 0.007. Similar values were also estimated for the broader Aegean region in previous studies (e.g. Karamanos et al., 2007; Mesimeri et al., 2018; Andinisari et al., 2020; Karakostas et al., 2021). Next, we proceeded to the determination of a crustal model using the VELEST algorithm (Kissling et al., 1994) by testing multiple published models as reference ones (e.g. Akyol et al., 2006; Karabulut et al., 2006; Konstantinou, 2018). All resulting models after multiple iterations consistently converged to a similar model (apart from the first few kilometers where differences could be found) very close to the model of Karabulut et al. (2006). We thus decided to adopt this model as a reference one (Figure S3) and use our derived model (Table 1) for the relocation procedure. We also added the appropriate station corrections calculated with the VELEST application to incorporate lateral inhomogeneities in the 1-D crustal model.

We then run the hypoDD program (Waldhauser, 2001) that employs the double difference algorithm (Waldhauser and Ellsworth, 2000), to improve the accuracy of the focal coordinates as they were obtained from HYPOINVERSE. Travel time differences between manually picked phases in the earthquake catalog were calculated, and then we kept event pairs with at least eight (8) observations, resulting to 32,956 P-phase pairs and 25,870 S-phase pairs (13 links per pair on average) in total. The differential times data set was analyzed using five sets with five iterations on each set gradually decreasing the residual threshold (secs) and the maximum distance (km) between catalog linked pairs. The horizontal and vertical errors between the initial and the relocated catalog indicate a substantial reduction of location uncertainty (Figure S4). The final catalog contains 817 relocated earthquakes out of 841 (~97%) that constituted the initial data set. The 477 earthquakes occurred before the 2022 main shock and the 340 are the aftershocks of the second sequence.

Depth (km)	V_p (km/s)
0.0–2.0	3.56
2.0–4.0	4.32
4.0–7.0	5.43
7.0–14.0	6.08
14.0–20.0	6.22
20.0–26.0	6.55
26.0–29.0	6.8
≥30	7.29

Table 1 P–wave velocity (V_p) model adjusted for the study area

2.2 Fault plane solutions

Fault plane solutions of 12 earthquakes with M≥4.0, which occurred during the whole study period, were calculated. One of them occurred slightly later, in April 2022, but in the same area. The moment tensor inversions were conducted using the Grond software (Heimann et al., 2018) which operates under the Pyrocko toolbox framework (Heimann et al., 2017). The methodology applied within Grond aims to minimize the misfit between synthetic and observed data by implementing a Bayesian bootstrapping inversion in parallel bootstrap chains. In the present case, the bootstrap algorithm was applied to minimize the L2 norm misfit for 22000 iterations, 3000 to uniformly sample the solution space and 19000 for the direct sampling. The number of parallel bootstrap chains was set equal to 200. We used the recordings of the regional broadband seismological stations of HUSN in distances ranging between 50 and 300 km (Figure S1). Green's functions were calculated for the local velocity model (Table 1) using the QSEIS program (Wang, 1999). A point source model was considered to perform the Bayesian optimization for a deviatoric moment tensor. Lastly, bandpass filters were applied, using frequencies between 0.04 and 0.11 Hz (the corresponding window for each earthquake is presented in Table 2) and the data fitting was carried out in the time domain. The NW-SE striking nodal planes of the focal mechanisms (Table 2; Table S1) exhibit predominantly sinistral strike-slip sense of motion, ranging from almost pure strike-slip to oblique faulting.

No	Date	Origin time	Lat (°)	Long (°)	Depth (km)	M_W	Freq (Hz)	Strike (°)	Dip (°)	Rake (°)	Misfit^2
1	2020/09/26	18:39:20	39.959	24.308	21.3	4.3	0.04–0.09	334	86	09	0.135
2	2020/09/26	22:50:25	39.954	24.306	18.9	5.3	0.04–0.08	145	73	-18	0.132
3	2020/09/27	12:22:20	39.983	24.315	11.5	4.4	0.04–0.09	131	51	-42	0.106
4	2020/09/27	12:58:02	39.943	24.320	15.6	4.3	0.04–0.09	339	85	-44	0.175
5	2020/09/27	16:05:08	39.952	24.329	19.1	4.1	0.05–0.10	310	53	-71	0.164
6	2020/09/28	04:12:42	39.943	24.296	18.9	4.7	0.04–0.08	147	76	-07	0.095
7	2022/01/16	11:48:06	39.980	24.292	17.6	5.4	0.04–0.08	125	52	-43	0.100
8	2022/01/16	12:26:19	39.999	24.276	12.7	4.5	0.05–0.09	185	68	18	0.109
9	2022/01/16	13:36:58	39.990	24.302	11.2	4.0	0.05–0.09	132	63	-33	0.144
10	2022/01/16	18:29:27	39.977	24.314	11.7	4.1	0.06–0.1	135	66	-28	0.111
11	2022/01/16	22:32:00	39.989	24.317	12.4	4.3	0.05–0.09	128	65	-31	0.091
12	2022/04/20	00:29:03	39.934	24.331	8.0	4.0	0.06–0.11	336	68	-30	0.124

Table 2 Fault plane solutions calculated for eleven earthquakes of the relocated catalog and one that occurred one month later (No 12). The focal parameters provided correspond to the relocated ones. In the last column, the square of the misfit of each solution is reported.

2.3 The Epidemic-Type Aftershock Sequence (ETAS) model

Short-term temporal properties of both the 2020 and 2022 seismic excitations were studied through the ETAS stochastic model (Ogata, 1988, 1998). The temporal ETAS model (Ogata, 1988) expresses the seismicity occurrence rate density as the summation of two parts, the constant background seismicity rate, μ, assumed time-independent, and the occurrence density rate of triggered earthquakes, $\lambda_i(t)$, and is given by a conditional intensity function, $\lambda(t)$:

$$\lambda(t) = \mu + \lambda_i(t) = \mu + \sum_{[i,t_i<t]} \frac{Ke^{\alpha(m_i - m_c)}}{(c + t - t_i)^p}, \qquad (1)$$

where K and α are the aftershock productivity parameters, c and p are the parameters of the temporal aftershock decay rate of the modified Omori law (Utsu, 1961), m_i is the magnitude of each earthquake occurred at time t_i and m_c is the completeness magnitude. The parameter K represents the intensity of aftershocks generation above m_c triggered by an earthquake with $m = m_c$, whereas α parameter describes the efficiency of earthquakes in triggering their own aftershocks. Large α values [1.3-3.1] indicate that large magnitude earthquakes trigger a large number of aftershocks, whereas small α [0.35-0.85] implies relatively higher triggering capabilities of small earthquakes. This means that Mainshock-Aftershock sequences typically tend to have larger α values, dominated by the main shock magnitude, while a swarm-like activity is characterized by small α values (Hainzl and Ogata, 2005; Ogata, 1992). The exponent p of the modified Omori law controls the aftershocks decay rate, where the decay is becoming faster as the value of p is increasing. The parameter c is linked with the short-term aftershock incompleteness soon after the occurrence of a main shock ($t = 0$), aiming to avoid singularities at occurrence times very close to $t = 0$. The five parameters of the model were estimated via the Maximum Likelihood Estimation (MLE) method using the simulated annealing technique proposed by Lombardi (2015) and implemented through an algorithm which is part of the SEDA package (Lombardi, 2017).

For the model evaluation, a residual analysis was performed Ogata (1988, 1998), which offers a qualitative model evaluation through visual display. Specifically, the occurrence times, t_i, were converted into transformed times, τ_i, according to the function:

$$\tau_i = \int_0^{t_i} \lambda(t)dt. \qquad (2)$$

Transformed times, τ_i, express the number of earthquakes that are expected to occur in the time interval [0, t_i]. If the estimated model adequately describes the temporal seismicity evolution, then transformed data, so-called residuals, behave like a stationary unit rate Poisson process. Otherwise, the transformed process will show some systematic departure from the linear Poisson process. Positive and negative departures indicate that the estimated model is under- and over-predicting the observed seismicity, respectively. In order to evaluate whether the transformed times, τ_i, are described by the Poisson process, the one-sample Kolmogorov-Smirnov test (KS1 test; Massey (1951)) was applied. More specifically, if τ_i are modelled by the unit rate Poisson process, then the transformed earthquake interevent times, $t_{i+1} - t_i$, should be independent and identically drawn from an exponential distribution (Llenos and Michael, 2013). The KS1 test is implemented under the null hypothesis that the earthquake interevent times are following the exponential distribution based on the p-value returned by the test, compared with the 0.05 significance level. If p-value is greater (or lower) than 0.05 then the null hypothesis can either be rejected or accepted.

3 Results

3.1 Spatiotemporal evolution of seismicity

The number of precisely relocated earthquakes equals to 817, with 477 of them belonging to the period starting from the initiation of the seismic excitation (26 September 2020) until the start of the second seismic sequence (16 January 2022) and the rest 340 constituting the second aftershock sequence. Regarding pre-seismic activity, only two (2) foreshocks were detected prior to the first main shock with the largest one occurring 4 hours in advance (Table 2). Our interpretations on aftershock activity and the associated active faults can be ensured by the quality criteria that the relocation procedure has fulfilled concerning the uncertainty estimation. These criteria do not prevent from systematic bias, but with our efforts to approach the velocity structure as best as possible we are confident that this bias could be considered small.

We are interested in how seismicity evolved spatially or temporally in the activated area on the temporal scale of several months, for which that this activity persisted. We started our relocated data set since 1 January 2020, almost 10 months before the first main shock occurrence. We may observe that seismicity is quite sparse with a couple of 2.0<M<2.9 shocks located close to the first main shock epicenter (green dots since the beginning of the space–time plot up to 26 September 2020, in Figure 2a) but not close in time for being considered foreshock activity as part of the nucleation phase. An M_w4.3 earthquake, instead, occurred about 4 hours before and in close distance with the main shock (Table 2). Intense aftershock activity followed the main shock on 26 September 2020, extended in an area more than 15 km long, much larger than the rupture length of an M_w5.3 main shock, (roughly 6 km) as prescribed by empirical scaling laws (Wells and Coppersmith, 1994; Thingbaijam et al., 2017).

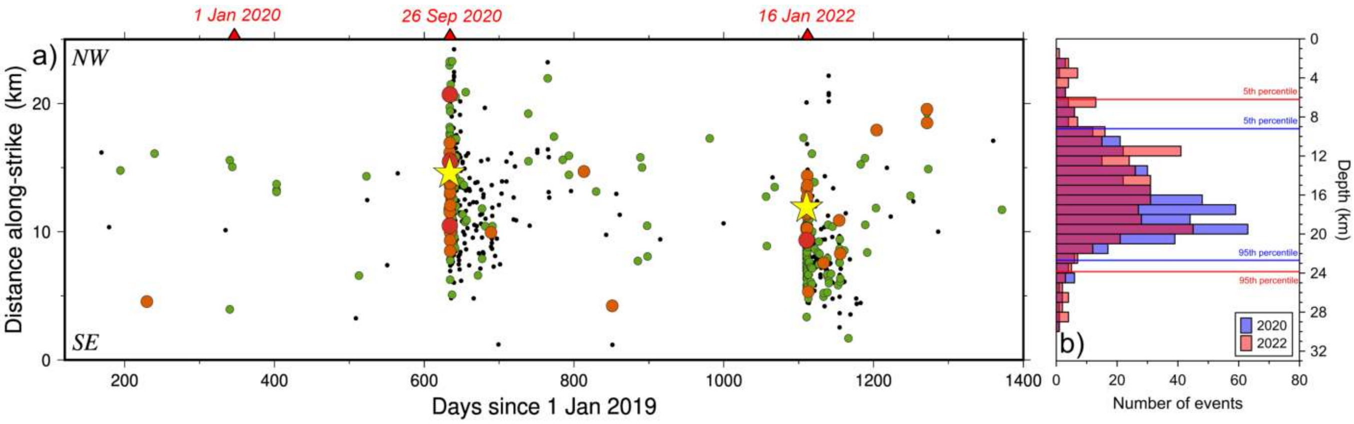


Figure 2 (a) Spatiotemporal distribution of the relocated seismicity since 1 January 2019 up to 1 November 2022. Yellow stars represent the two main shocks, red circles the $4.0 \leq M_w \leq 4.9$, orange circles the $3.0 \leq M \leq 3.9$, green circles the $2.0 \leq M \leq 2.9$, and small black dots the $M \leq 2.0$ earthquakes. (b) Histogram of the focal depths with different colors denoting the two aftershock sets.

As Figures 2a & 3a show, seismicity is mainly concentrated on a 20 km length prolonged zone, with a general strike of 145° (thick dashed line in Figure 3a). The epicentral distribution of the two distinct clusters present a spatial overlap of slightly over 50%. The determined fault plane solutions unveil the dominance of strike-slip faulting in the study area given that pure strike-slip (e.g., No 1 & 6, Table 2; Figure 3a) and strike-slip with a normal component (No 2 & 4 & 9-12) focal mechanisms are present. The complex faulting patterns coming into play are revealed by the presence of either pure normal (No 5, Table 2; Figure 3a) or oblique (normal with a strike-slip component) focal mechanisms (No 3 & 7, Table 2; Figure 3a).

In Figures 4 & 5 the relocated seismicity of the two activated structures is shown separately aiming to detail the properties of each seismic sequence. Regarding the 2020 M_w5.3 sequence (Figure 4a), the epicentral alignment (NW-SE) of the aftershocks is in good agreement with the strike of the one nodal plane of the main shock focal mechanism. The total length of the activated area is approximately 15 km, which significantly exceeds the estimations of frequently used empirical relationships between main shock magnitude and the causative fault length (Wells and Coppersmith, 1994; Thingbaijam et al., 2017). The duration of the intense aftershock activity was short, with over half of the aftershocks occurring within the first week after the main shock (Figure 3). Moreover, all M≥4.0 aftershocks took place in less than 3 days (Table 2). An interesting feature of this seismic sequence is also the occurrence of a few foreshocks with the strongest among them having a magnitude of M_w4.4 (Table 2), placed very close to the main shock epicenter (Figure 2) although in deeper parts of the seismogenic layer.

The aftershock activity during the first 2 days defines a seismogenic zone of 9 km (13 – 22 km) (Figure 4b). As time advances, however, the inclusion of all earthquakes (Figure 4c) evidences an expansion of the aftershock zone in more shallow depths (~5km). Normal to the strike cross section encompassing events within the larger part of the aftershock distribution (3 km of either side of the cross section) provides a complete picture of the activated fault plane (Figure 4d). The total depth extent of the aftershock zone ranges from 6 to 24 km (Figure 2b).

Shifting our focus to the second seismic sequence that emerged from the strongest event included in our catalogue (M_w5.4) at the beginning of 2022 we might observe that the epicentral alignment of the aftershocks is analogous to the previous case (NW-SE) however their spatial extent is developed further to the north. The total length of the activated area, as inferred by the aftershock epicentral distribution is approximately equal to 10 km (Figures 2a, 5a). The aftershock activity of the first 2 days occupies a depth range of 9 to 21 km, inside a 12 km extent (Figure 5b), with only minimal low-magnitude events outlying. Figure 5c, which includes all events at a distance of 1.5 km either side of the normal cross section, appears to be very similar to the one of Figure 4b, highlighting the fast decay of the aftershock rate after the first few days. Moreover, the wider

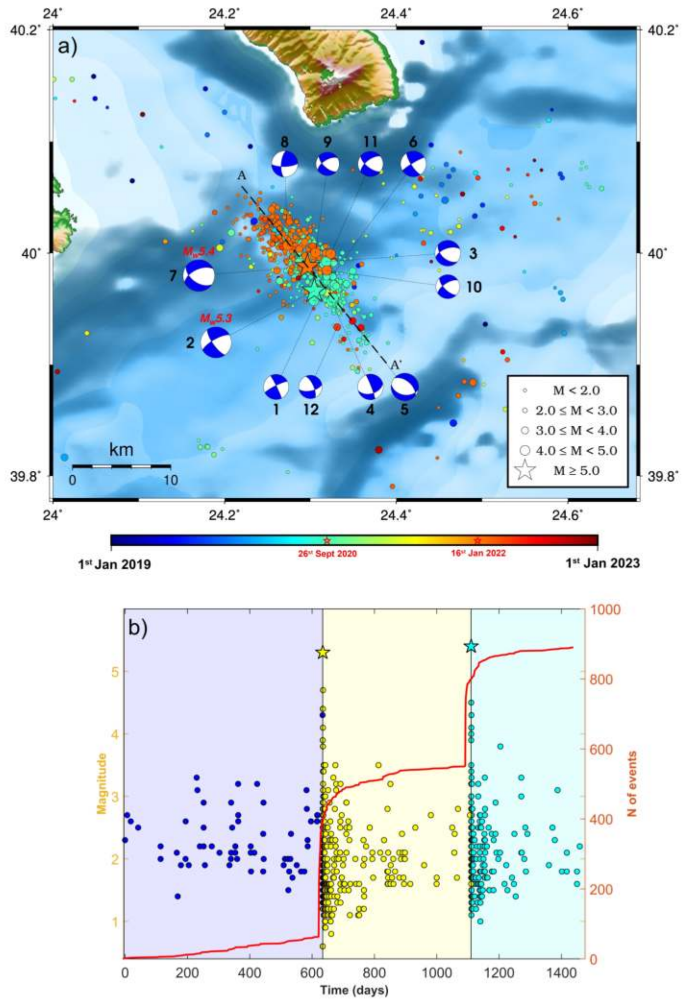


Figure 3 Seismicity of the study area, since the beginning of 2019 until the end of 2022. (a) Stars indicate the relocated epicenters of the main shocks, whereas circles denote either the relocated aftershock locations or the background seismicity. All epicenters are scaled in compliance with the corresponding earthquake magnitudes and color-coded according to the temporal scale at the bottom of the figure. Fault plane solutions determined in this study (Table 2) are also shown as lower hemisphere equal area projections, with the compressional quadrants colored in blue. (b) Temporal evolution of seismicity for the 2019–2022 period versus magnitude. Blue, yellow and cyan circles and shaded areas indicate the pre–2020, 2020 and 2022 seismic activity. Red solid line depicts the cumulative number of earthquakes for the same period.

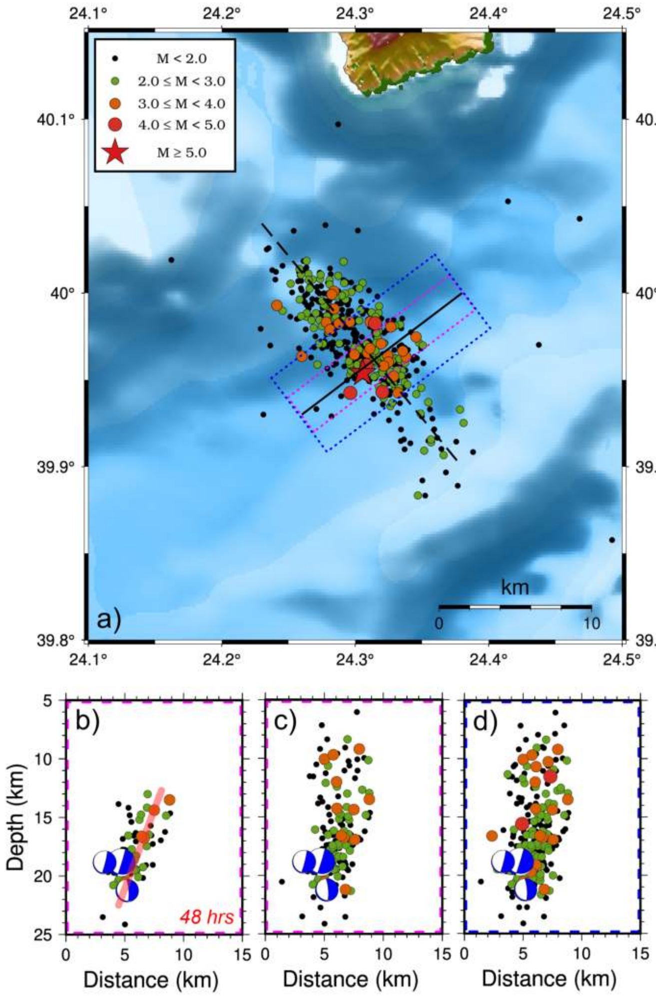


Figure 4 a) Spatial distribution of the relocated aftershocks of the M_w5.3 main shock on 26 September 2020 (red star). The epicenters of the aftershocks are shown as circles of different sizes and colors according to their magnitude. The black dashed line shows the general strike of the epicentral distribution, whereas the constant one signifies the normal to the strike cross sections including events b) within 1.5 km either side of the section, during the first 2 days after the main shock, c) within 1.5 km for the entire duration and d) 3.0 km either side of the section, again for the entire relocated catalog.

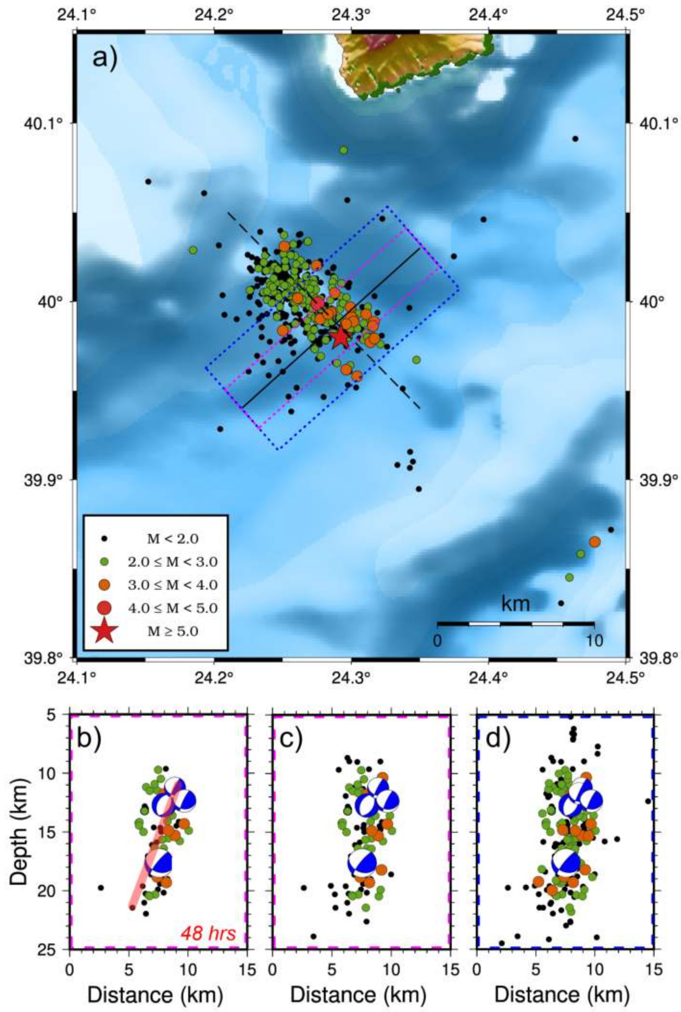


Figure 5 a) Spatial distribution of the relocated aftershocks of the M_w5.4 main shock on 16^{th} January 2022 (red star). The epicenters of the aftershocks are plotted with circles of different sizes and colors according to their magnitude. The black dashed line shows the general strike of the epicentral distribution, whereas the constant one signifies the normal to the strike cross sections including events b) within 1.5 km either side of the section, during the first 2 days after the main shock, c) within 1.5 km for the entire duration and d) 3.0 km either side of the section, again for the entire relocated catalog.

normal cross section (Figure 5d) reveals the close proximity of the strongest aftershocks (11-13 km in depth).

3.2 Temporal features from the ETAS model fitting

The temporal ETAS parameters (μ, K, α, c and p) were estimated through the MLE method after the determination of the magnitude of completeness, m_c, for the entire initial earthquake catalog covering the period 2019-2023. The m_c was identified through the Goodness-of-Fit method (GFT; Wiemer and Wyss (2000)), considering the 95% confidence level of residuals (Figure S5) and was found equal to m_c=1.9 (residuals value equal to 3.66%) resulting to a data set of 470 earthquakes with $m \geq m_c$ and a b-value equal to 0.78 (b=0.78). The model was first applied in the entire initial earthquake cata-

Period	μ	K	α	c	p	Obs.	KS1 p-value
2019 −2022	0.046	0.015	1.78	0.030	1.23	470	0.62
01/01/2020 −30/6/2021	0.047	0.017	1.75	0.029	1.17	240	0.60
1/7/2021 −31/12/2022	0.037	0.011	1.86	0.037	1.25	197	0.57

Table 3 Temporal ETAS parameters estimates (μ, K, α, c and p) for the periods 2019-2022, January 2020-June 2021 and July 2021-December 2022, along with the respective number of observations and the p-values of the one sample Kolmogorov-Smirnov goodness-of-fit test (KS1) between the earthquake interevent times and the exponential distribution.

log of the period 2019-2022 and then in two additional and distinctive sub-periods, namely from January 2020 to June 2021 and from July 2021 to December 2022, aiming to compare the temporal properties of the two seismic sequences. The selection of each sub-period was based on the adequacy of the number of earthquakes before the occurrence of each excitation, in which the parameters were estimated, representing the learning phase of the ETAS model application. The calculated parameter values are given in Table 3.

The ETAS application for the entire period (2019-2022) ascertained that the model adjusts well the observed earthquake rate, with slight discrepancies soon after the occurrence of the 2020 M_w5.3 and 2022 M_w=5.4 main shocks (Figure 6a). The good fit of the estimated model in respect to the observations is highlighted by the residuals analysis application (Figure 6b), since the lines that depict the expected and the observed number of events against the transformed times (red and black lines in Figure 6b, respectively) almost coincide. This latter fact is confirmed by the result of the KS1 goodness of fit test shown in Table 3. Specifically, the calculated p-value of the test is equal to 0.62, much larger than the 0.05 confidence level. The estimated parameters are acquiring values typical for Mainshock-Aftershock sequences (Ogata, 1992). In more detail, the observed seismicity during the period from 2019 to 2022 is characterized by a very low background rate equal to μ=0.046 event/day. This means that almost 88% of the total number of earthquakes are offsprings of the 2020 M_w5.3 and 2022 M_w=5.4 sequences, and only 68 out of the 470 are assumed to be independent. Furthermore, the productivity parameter, α, was estimated equal to $\alpha = 1.78$, indicating also Mainshock-Aftershock type of sequence, since Mainshock-Aftershock activity is typically characterized by α values ranging between [1.3-3.1], whereas swarm-like activity is ascribed to values ranging between [0.35- 0.85] (Hainzl and Ogata, 2005; Ogata, 1992). The parameters expressing temporal characteristics, c and p, attain expected values, in comparison with those reported in previous studies ranging from 0.03 to 0.07 and from 1.16 to 1.25, respectively (Chu et al., 2011; Kourouklas et al., 2020).

Placing emphasis on the two sub-periods (January 2020-June 2021 and July 2021-December 2022), when the M_w5.3 and M_w5.4 main shocks occurred, the small discrepancies between the observed earthquake rates and the modeled ones are becoming more explicit (Figures 6c and 6e, for the January 2020-June 2021 and July 2021-December 2022, respectively). Specifically, it is observed that the estimated models for both periods slightly underestimate the earthquake rates soon after the occurrence of the M_w5.3 and M_w5.4 earthquakes. This result is also visible from the residual analysis plots (Figures 6d & 6f), in which the observed transformed times (red lines in Figures 6d & 6f) slightly deviate from the unit rate Poisson process (black lines in Figures 6d & 6f). However, the p-values of the KS1 test (p-value=0.60 and 0.57 for the periods January 2020-June 2021 and July 2021-December 2022, respectively; Table 3) are again much larger than the significance level (0.05), suggesting a good performance of the temporal ETAS model to the data from a statistical point of view.

Comparison of the parameter estimates of the two sub-periods (Table 3) shows that throughout the first one (January 2020-June 2021), during which the M_w5.3 earthquake occurred (September 26th of 2020), the background rate is estimated equal to $\mu = 0.047$ event/day, a value almost equal to the entire period estimate). On the contrary, background rate is found quite smaller ($\mu = 0.037$ event/day) during the period from July 2021 to December 2022, in which the M_w5.4 earthquake occurred (January 16th of 2022). Both estimated values are again indicating a clear Mainshock-Aftershock activity as well as the estimated values of the productivity parameter, α, ($\alpha = 1.75$ and 1.86 for the 1st and the 2nd sub-periods, respectively; Table 3) with the one attached to the second sub-period being higher. Additionally, the estimated value of p parameter of the modified Omori law is larger in the second sub-period's application ($p = 1.17$ instead of $p = 1.25$ for the 1st period) indicating the smaller duration of the second sequence.

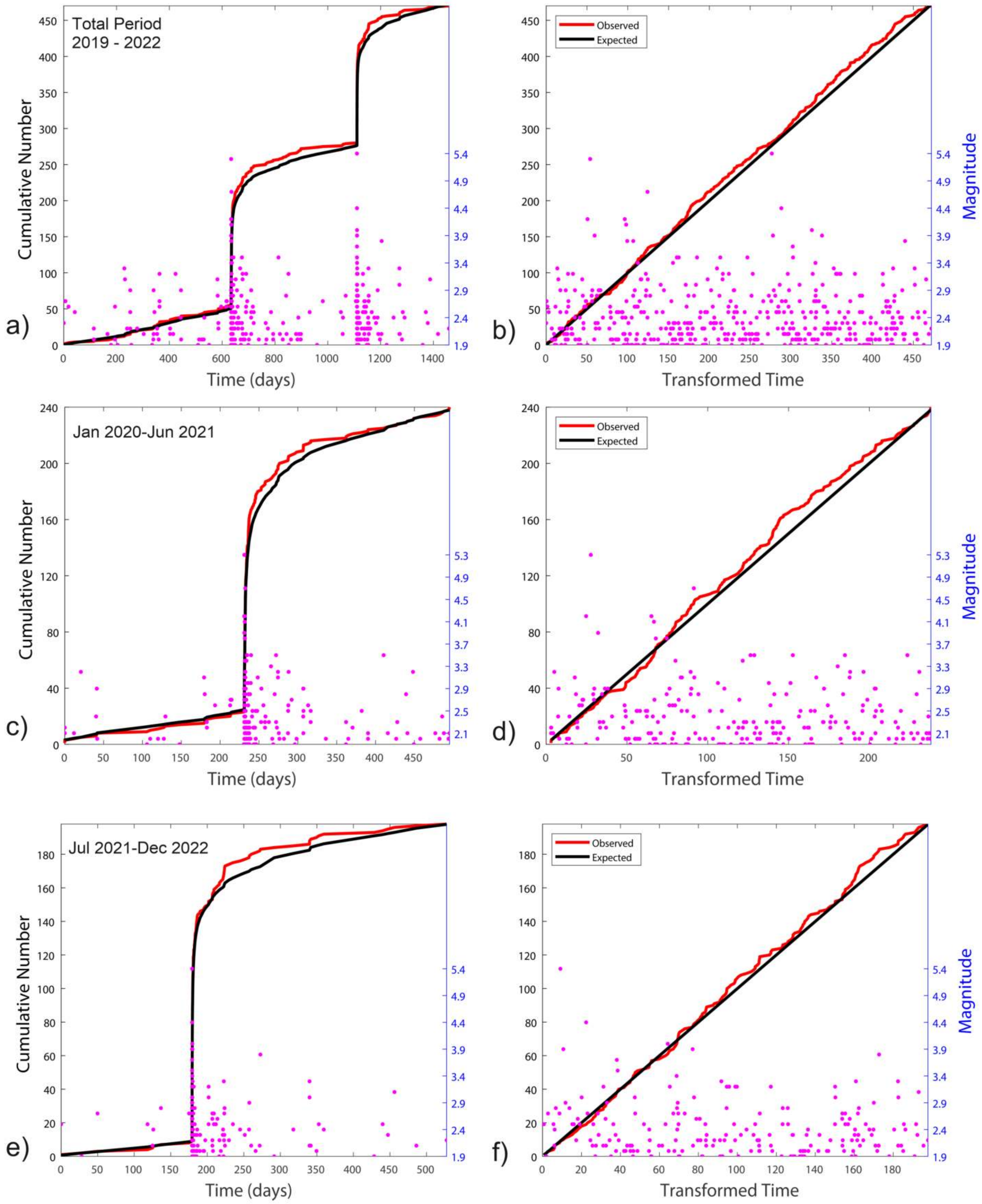


Figure 6 Observed (black lines) and expected by the estimated temporal ETAS model (red lines) cumulative earthquake number against ordinary (a,c,e) and transformed (b,d,f) time for the periods 2019 – 2022 (a & b, respectively), January 2020 – June 2021 (c & d, respectively) and July 2021 – December 2022 (e & f, respectively). Magnitudes of earthquakes with $m \geq m_c$ (right y-axis) against time are shown as pink dots in all subplots.

4 Discussion

As the westward prolongation of the North Anatolian Fault (NAF) system into the Aegean Sea, the North Aegean Trough (NAT) is characterized by dextral strike-slip faulting. The intense N-S extension of the back-arc area due to the roll back of the subducted slab in southern Aegean, resulted to a complex transtensional basin dated back to early Pliocene or Pleistocene. The conjunction of major strike-slip systems with a wide variety of secondary structures, especially at fault zone tips or between linked structures, the damage zones (e.g Petit and Barquins, 1988; Kim et al., 2004; Kim and Sanderson, 2006) is very commonly met. The reason behind their development can be attributed to stress concentrations (e.g Cox and Scholz, 1988) or to host displacement alterations along the fault zones (Kim et al., 2000). The North Aegean area is characterized by a broad spectrum of strike-slip damage patterns (Figure 7), with the most important and relevant to our study area being:

a) Horsetail structures: The western endpoint of the NAT is characterized by a right-stepping horsetail structure (red patch in Figure 7) expressed through numerous oblique splays stemming from the main strike-slip zone forming a number of transtensive basins (Sakellariou et al., 2017).

b) Positive or negative flower structures: Negative flower structures are found inside most of the sub-basins bounded by the oblique splays. They consist of opposing-dip normal fault structures and merge into a single strike-slip fault at deeper parts (upper left part rectangle in Figure 7). Positive ones are also present but outside the area under study, along the south Marmara and the Ganos fault segments, expressing transpressive structures owing their formation to strike-slip faults with a reverse component (Rodriguez et al., 2023).

c) Conjugate faults: They are faults that intersect with the main structure at a high angle (typically more than 60°) and exhibit a sense of displacement opposite to the dominant one, with which they are often not connected. In our case, they represent the sinistral counterparts of the prevalent dextral strike-slip faults that dominate the northern Aegean area. Conjugate faults may be combined with branch faults (see below) and create block rotation (e.g. Nicholson et al., 1986; Kim et al., 2003).

d) Branch faults: They represent shear fractures having similar sense of motion as the main strike-slip zone (dextral strike-slip in our case, blue patch in Figure 7). They can act together with the main zone to create splays (e.g. Kim et al., 2004) or with other structures and form more complex patterns.

The diversity characterizing the termination of strike-slip fault systems, especially on such a large scale such as the NAT makes it very challenging to uncover its fine details, due to their complex nature and high intercorrelation. Typical examples are the fault segments associated with the two main shocks of this study which are not included in the simplified tectonic map of the North Aegean based on the interpretation of the air gun lithospheric profiles by Papanikolaou et al. (2006) as well as in other studies (e.g. Ferentinos et al., 2018; Sakellariou and Tsampouraki-Kraounaki, 2019).

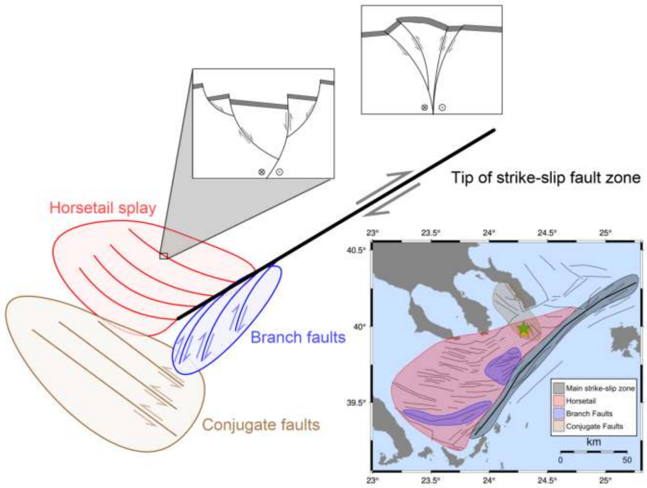


Figure 7 Strike-slip damage patterns (schematic illustrations) and their connection to the study area (simplified map). The green and yellow stars indicate the epicenters of the main shocks of this study.

Based on our analysis, the activated structures hosting the 2020 M_w5.3 and 2022 M_w 5.4 main shocks can be attributed to conjugate faults at the termination of the main strike-slip zone, having a considerable vertical component consistent with the extensional character of the NAF termination (Figure 7). The primary indications leading to this suggestion are the determined focal mechanisms of the main shocks and the stronger aftershocks (Table 2), which indicate an intricate stress field consisting of strike-slip and normal faulting style. Adding to this, the strike-slip moment tensor solutions exhibit a left-lateral displacement, as opposed to the dextral strike-slip motion characterizing the NAF. Moreover, the spatiotemporal distribution of aftershocks (Figures 3-5) further demonstrates the high angle (almost perpendicular) in which the activated structures lie in comparison to the NAF (Figure 1). The scenario of a transtensional flower structure cannot be ruled out given that such systems are closely related and common to conjugate fault systems. A valid reason for this argument can be ascribed to the relative differences of the focal depths of the two sequences (Figure 2). Even though the spatial distribution of both aftershock sequences appears to extend in more or less the same area in map view (Figure 3), the mean focal depths from the aftershocks originated from the 2020 M_w5.3 sinistral strike-slip main shock are found in deeper parts compared to those from the 2022 M_w5.4 oblique (λ=-43°; normal with sinistral strike-slip component) one. In this case, a depressed area, or even a pull-apart basin (on a larger scale) is most commonly formed. Bathymetric maps (Papanikolaou et al. (2006) and references therein) cannot provide us with a definite answer however our study area is located at the northern margins of the main North Aegean basin and more specifically in an area where a bulge is disrupting the almost linear NE-SW oriented edges of the basin.

Statistical analysis of the short-term clustering features of the 2020-2022 excitations highlights their

Mainshock-Aftershock nature. The estimated temporal ETAS parameters show non-significant background seismicity rate changes for both the average model (2019-2022) and the two sub-period models, which are referring to the 2020 and 2022 sequences, indicating that the excitation is driven by the regional tectonic loading. An interesting remark arises from the estimated parameters of the second sequence (2022) modeling, in which the temporal parameters (c and p) are larger. These values indicate a faster decay rate compared to the first sequence's aftershocks, even though the magnitude of the 2022 main shock is larger than the 2020 one. This remark could be likely explained by the fact that the study area was recently activated and consequently rather stress relaxed.

Moreover, the low background rate, the estimated productivity parameters and the fast aftershock decay rates in all evaluated models provide an indirect insight into the properties of the activated fault segments, qualifying them as rather weak. This property characterizing fault segments in the broader area (including the main branch of NAT fault zone) has already been suggested by physics-based earthquake simulator results (Kourouklas et al., 2021).

5 Conclusions

We thoroughly investigated the seismic sequences of two moderate main shocks occurred in the region north of NAT on 26 September 2020 (M_w5.3) and 16 January 2022 (M_w5.4), by means of their spatiotemporal distribution, moment tensor solutions and statistical analysis. We constrained the dimensions of the rupture areas (10x10 km^2 for the 2020 main shock and 11x10 km^2 for the 2022 main shock) based on the relocated early aftershocks. Fault plane solutions highlight the complexity of the faulting patterns, with most of them exhibiting mainly sinistral strike-slip faulting, with some oblique cases (strike-slip with a normal component) being also present.

All things considered, we attribute the occurrence of the investigated seismic sequences as a result of secondary faulting related to the termination of an active strike-slip plate boundary, expressed as transtensional deformation. We may consider these main shocks as a beneficial example of better recognition of the complex behavior characterizing the NAT, but through the lens of seismic hazard, we weigh this kind of events much less prominent than the overall seismic hazard of the broader area.

6 Acknowledgements

We would like to thank Kostas Leptokaropoulos and an anonymous reviewer for their valuable comments that helped to improve this manuscript. We also thank Yen Joe Tan for his editorial assistance. The SEDA software package (Lombardi, 2017) was used for the estimation of the temporal ETAS parameters (Available online: https://zenodo.org/record/55277; last accessed on March 2023). The software package of Pyrocko Heimann et al. (2017) used for the moment tensor inversions is available at http://pyrocko.org. Figures were created with the use of MATLAB software (Available online: http://mathworks.com; last accessed in March 2023) and Grapher version 16 (http://www.GoldenSoftware.com). The software Generic Mapping Tools (Wessel et al., 2013) was used to plot the maps and the cross sections. Geophysics Department Contribution 971.

7 Data and Code Availability

The seismic data used in this study are publicly available at https://doi.org/10.7914/SN/HT.

References

Akyol, N., Zhu, L., Mitchell, B. J., Sözbilir, H., and Kekovalı, K. Crustal structure and local seismicity in western Anatolia. *Geophysical Journal International*, 166(3):1259–1269, Sept. 2006. doi: 10.1111/j.1365-246x.2006.03053.x.

Andinisari, R., Konstantinou, K., and Ranjan, P. Seismotectonics of SE Aegean inferred from precise relative locations of shallow crustal earthquakes. *Journal of Seismology*, 24(1):1–22, 2020.

Bitharis, S., Pikridas, C., Fotiou, A., and Rossikopoulos, D. GPS data analysis and geodetic velocity field investigation in Greece, 2001–2016. *GPS Solutions*, 28(1), Oct. 2023. doi: 10.1007/s10291-023-01549-8.

Chu, A., Schoenberg, F. P., Bird, P., Jackson, D. D., and Kagan, Y. Y. Comparison of ETAS Parameter Estimates across Different Global Tectonic Zones. *Bulletin of the Seismological Society of America*, 101(5):2323–2339, Sept. 2011. doi: 10.1785/0120100115.

Cox, S. and Scholz, C. On the formation and growth of faults: an experimental study. *Journal of Structural Geology*, 10(4):413–430, Jan. 1988. doi: 10.1016/0191-8141(88)90019-3.

Drakopoulos, J. C. and Ekonomides, A. C. Aftershocks of February 19, 1968 earthquake in Northern Aegean Sea and related problems. *Pure and Applied Geophysics PAGEOPH*, 95(1):100–115, 1972. doi: 10.1007/bf00878858.

England, P., Houseman, G., and Nocquet, J. Constraints from GPS measurements on the dynamics of deformation in Anatolia and the Aegean. *Journal of Geophysical Research: Solid Earth*, 121 (12):8888–8916, Dec. 2016. doi: 10.1002/2016jb013382.

Ferentinos, G., Georgiou, N., Christodoulou, D., Geraga, M., and Papatheodorou, G. Propagation and termination of a strike slip fault in an extensional domain: The westward growth of the North Anatolian Fault into the Aegean Sea. *Tectonophysics*, 745: 183–195, Oct. 2018. doi: 10.1016/j.tecto.2018.08.003.

Hainzl, S. and Ogata, Y. Detecting fluid signals in seismicity data through statistical earthquake modeling. *Journal of Geophysical Research: Solid Earth*, 110(B5), Mar. 2005. doi: 10.1029/2004jb003247.

Hatzfeld, D., Ziazia, M., Kementzetzidou, D., Hatzidimitriou, P., Panagiotopoulos, D., Makropoulos, K., Papadimitriou, P., and Deschamps, A. Microseismicity and focal mechanisms at the western termination of the North Anatolian Fault and their implications for continental tectonics. *Geophysical Journal International*, 137(3):891–908, June 1999. doi: 10.1046/j.1365-246x.1999.00851.x.

Heimann, S., Kriegerowski, M., Isken, M., Cesca, S., Daout, S., Grigoli, F., Juretzek, C., Megies, T., Nooshiri, N., Steinberg, A., Sudhaus, H., Vasyura-Bathke, H., Willey, T., and Dahm, T.

Pyrocko-An open-source seismology toolbox and library [Computer software. GFZ Data Services, 2017.

Heimann, S., Isken, M., Kuhn, D., Sudhaus, H., Steinberg, A., Daout, S., Cesca, S., Vasyura-Bathke, H., and Dahm, T. *Grond—A probabilistic earthquake source inversion framework V1.0 [Computer software*. GFZ Data Services, 2018.

Kapetanidis, V. and Kassaras, I. Contemporary crustal stress of the Greek region deduced from earthquake focal mechanisms. *Journal of Geodynamics*, 123:55–82, Jan. 2019. doi: 10.1016/j.jog.2018.11.004.

Karabulut, H., Roumelioti, Z., Benetatos, C., Mutlu, A. K., Özalaybey, S., Aktar, M., and Kiratzi, A. A source study of the 6 July 2003 (Mw 5.7) earthquake sequence in the Gulf of Saros (Northern Aegean Sea): Seismological evidence for the western continuation of the Ganos fault. *Tectonophysics*, 412(3–4):195–216, Jan. 2006. doi: 10.1016/j.tecto.2005.09.009.

Karakostas, B., Hatzidimitriou, P., Karakaisis, G., Papadimitriou, E., and Papazachos, B. Evidence for long–term precursors of strong earthquakes in the northernmost part of the Aegean Sea. *Earthquake Prediction Research*, 4:155–164, 1986.

Karakostas, V., Papadimitriou, E., and Gospodinov, D. Modelling the 2013 North Aegean (Greece) seismic sequence: geometrical and frictional constraints, and aftershock probabilities. *Geophysical Journal International*, 197(1):525–541, Jan. 2014. doi: 10.1093/gji/ggt523.

Karakostas, V., Tan, O., Kostoglou, A., Papadimitriou, E., and Bonatis, P. Seismotectonic implications of the 2020 Samos, Greece, Mw 7.0 mainshock based on high-resolution aftershock relocation and source slip model. *Acta Geophysica*, 69(3):979–996, Apr. 2021. doi: 10.1007/s11600-021-00580-y.

Karakostas, V. G., Papadimitriou, E. E., Karakaisis, G. F., Papazachos, C. B., Scordilis, E. M., Vargemezis, G., and Aidona, E. The 2001 Skyros, Northern Aegean, Greece, earthquake sequence: Off—Fault aftershocks, tectonic implications, and seismicity triggering. *Geophysical Research Letters*, 30(1), Jan. 2003. doi: 10.1029/2002gl015814.

Karamanos, C. K., Karakostas, G. V., Papadimitriou, E. E., and Sachpazi, M. Earthquake relocation in the wester termination of the North Aegean Trough. *Bulletin of the Geological Society of Greece*, 40(3):1091, June 2007. doi: 10.12681/bgsg.16828.

Kim, Y. and Sanderson, D. J. Structural similarity and variety at the tips in a wide range of strike–slip faults: a review. *Terra Nova*, 18 (5):330–344, Sept. 2006. doi: 10.1111/j.1365-3121.2006.00697.x.

Kim, Y.-S., Andrews, J. R., and Sanderson, D. J. Damage zones around strike-slip fault systems and strike-slip fault evolution, Crackington Haven, southwest England. *Geosciences Journal*, 4 (2):53–72, June 2000. doi: 10.1007/bf02910127.

Kim, Y.-S., Peacock, D., and Sanderson, D. J. Mesoscale strike-slip faults and damage zones at Marsalforn, Gozo Island, Malta. *Journal of Structural Geology*, 25(5):793–812, May 2003. doi: 10.1016/s0191-8141(02)00200-6.

Kim, Y.-S., Peacock, D. C., and Sanderson, D. J. Fault damage zones. *Journal of Structural Geology*, 26(3):503–517, Mar. 2004. doi: 10.1016/j.jsg.2003.08.002.

Kiratzi, A., Tsakiroudi, E., Benetatos, C., and Karakaisis, G. The 24 May 2014 (Mw6.8) earthquake (North Aegean Trough): Spatiotemporal evolution, source and slip model from teleseismic data. *Physics and Chemistry of the Earth, Parts A/B/C*, 95:85–100, Oct. 2016. doi: 10.1016/j.pce.2016.08.003.

Kiratzi, A. A., Wagner, G. S., and Langston, C. A. Source parameters of some large earthquakes in Northern Aegean determined by body waveform inversion. *Pure and Applied Geophysics PAGEOPH*, 135(4):515–527, Apr. 1991. doi: 10.1007/bf01772403.

Kissling, E., Ellsworth, W. L., Eberhart-Phillips, D., and Kradolfer, U. Initial reference models in local earthquake tomography. *Journal of Geophysical Research: Solid Earth*, 99(B10):19635–19646, Oct. 1994. doi: 10.1029/93jb03138.

Klein, F. W. *User's guide to HYPOINVERSE-2000, a Fortran program to solve for earthquake locations and magnitudes*. 2002. doi: 10.3133/ofr02171.

Konstantinou, K. Accurate relocation of seismicity along the North Aegean Trough and its relation to active tectonics. *Tectonophysics*, 717:372–382, Oct. 2017. doi: 10.1016/j.tecto.2017.08.021.

Konstantinou, K. Estimation of optimum velocity model and precise earthquake locations in NE Aegean: Implications for seismotectonics and seismic hazard. *Journal of Geodynamics*, 121: 143–154, Nov. 2018. doi: 10.1016/j.jog.2018.07.005.

Konstantinou, K. I., Mouslopoulou, V., Liang, W., Heidbach, O., Oncken, O., and Suppe, J. Present-day crustal stress field in Greece inferred from regional-scale damped inversion of earthquake focal mechanisms. *Journal of Geophysical Research: Solid Earth*, 122(1):506–523, Jan. 2017. doi: 10.1002/2016jb013272.

Kourouklas, C., Papadimitriou, E., Tsaklidis, G., and Karakostas, V. Earthquake recurrence models and occurrence probabilities of strong earthquakes in the North Aegean Trough (Greece). *Journal of Seismology*, 22(5):1225–1246, June 2018. doi: 10.1007/s10950-018-9763-8.

Kourouklas, C., Mangira, O., Iliopoulos, A., Chorozoglou, D., and Papadimitriou, E. A study of short-term spatiotemporal clustering features of Greek seismicity. *Journal of Seismology*, 24(3): 459–477, May 2020. doi: 10.1007/s10950-020-09928-1.

Kourouklas, C., Console, R., Papadimitriou, E., Murru, M., and Karakostas, V. Modelling the large earthquakes recurrence times along the North Aegean Trough Fault Zone (Greece) with a physics-based simulator, Geophys. *Geophysical Journal International*, 225(3):2135–2156, Mar. 2021. doi: 10.1093/gji/ggab085.

Kourouklas, C., Tsaklidis, G., Papadimitriou, E., and Karakostas, V. Analyzing the Correlations and the Statistical Distribution of Moderate to Large Earthquakes Interevent Times in Greece. *Applied Sciences*, 12(14):7041, July 2022. doi: 10.3390/app12147041.

Llenos, A. L. and Michael, A. J. Modeling Earthquake Rate Changes in Oklahoma and Arkansas: Possible Signatures of Induced Seismicity. *Bulletin of the Seismological Society of America*, 103 (5):2850–2861, Sept. 2013. doi: 10.1785/0120130017.

Lombardi, A. M. Estimation of the parameters of ETAS models by Simulated Annealing. *Scientific Reports*, 5(1), Feb. 2015. doi: 10.1038/srep08417.

Lombardi, A. M. SEDA: A software package for the Statistical Earthquake Data Analysis. *Scientific Reports*, 7(1), Mar. 2017. doi: 10.1038/srep44171.

Massey, F. J. The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association*, 46(253):68–78, Mar. 1951. doi: 10.1080/01621459.1951.10500769.

McKenzie, D. Active Tectonics of the Mediterranean Region. *Geophysical Journal International*, 30(2):109–185, Dec. 1972. doi: 10.1111/j.1365-246x.1972.tb02351.x.

Mesimeri, M., Kourouklas, C., Papadimitriou, E., Karakostas, V., and Kementzetzidou, D. Analysis of microseismicity associated with the 2017 seismic swarm near the Aegean coast of NW Turkey. *Acta Geophysica*, 66(4):479–495, May 2018. doi: 10.1007/s11600-018-0157-7.

Nicholson, C., Seeber, L., Williams, P., and Sykes, L. R. Seismic evidence for conjugate slip and block rotation within the

San Andreas Fault System, southern California. *Tectonics*, 5(4): 629–648, Aug. 1986. doi: 10.1029/tc005i004p00629.

Ogata, Y. Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes. *Journal of the American Statistical Association*, 83(401):9–27, Mar. 1988. doi: 10.1080/01621459.1988.10478560.

Ogata, Y. Detection of precursory relative quiescence before great earthquakes through a statistical model. *Journal of Geophysical Research: Solid Earth*, 97(B13):19845–19871, Dec. 1992. doi: 10.1029/92jb00708.

Ogata, Y. Space-Time Point-Process Models for Earthquake Occurrences. *Annals of the Institute of Statistical Mathematics*, 50(2): 379–402, June 1998. doi: 10.1023/a:1003403601725.

Pacheco, J. F. and Sykes, L. R. Seismic moment catalog of large shallow earthquakes, 1900 to 1989. *Bulletin of the Seismological Society of America*, 82(3):1306–1349, June 1992. doi: 10.1785/bssa0820031306.

Papanikolaou, D., Alexandri, M., and Nomikou, P. *Active faulting in the north Aegean basin*. Geological Society of America, 2006. doi: 10.1130/2006.2409(11).

Papazachos, B. and Papazachou, C. *The earthquakes of Greece*. Ziti Publ. Co, 2003.

Papazachos, B., Papadimitriou, E., Kiratzi, A., Papazachos, C., and Louvari, E. Fault plane solutions in the Aegean Sea and the surrounding area and their tectonic implication. *Bollettino Di Geofisica Teorica Ed Applicata*, 39:199–218, 1998.

Papazachos, B. C. and Comninakis, P. E. Geophysical and tectonic features of the Aegean Arc. *Journal of Geophysical Research*, 76 (35):8517–8533, Dec. 1971. doi: 10.1029/jb076i035p08517.

Papazachos, C. B. and Kiratzi, A. A. A detailed study of the active crustal deformation in the Aegean and surrounding area. *Tectonophysics*, 253(1–2):129–153, Mar. 1996. doi: 10.1016/0040-1951(95)00047-x.

Petit, J. and Barquins, M. Can natural faults propagate under Mode II conditions? *Tectonics*, 7(6):1243–1256, Dec. 1988. doi: 10.1029/tc007i006p01243.

Pichon, X. L. and Angelier, J. The hellenic arc and trench system: A key to the neotectonic evolution of the eastern mediterranean area. *Tectonophysics*, 60(1–2):1–42, Nov. 1979. doi: 10.1016/0040-1951(79)90131-8.

Rocca, A., Karakaisis, G., Karacostas, B., Kiratzi, A., Scordilis, E., and Papazachos, B. Further Evidence on the Strike-slip Faulting of the Northern Aegean Trough Based on Properties of the August–November 1983 Seismic Sequence. *Bolletino Di Geofisica Teorica Ed Applicata*, XXVII(106):101–109, 1985.

Rodriguez, M., Sakellariou, D., Gorini, C., Janin, A., D'Acremont, E., Pourhiet, L. L., Chamot-Rooke, N., Tsampouraki-Kraounaki, K., Morfis, I., Rousakis, G., Henry, P., Lurin, A., Delescluse, M., Briole, P., Rigo, A., Arsenikos, S., Bulois, C., Fernández-Blanco, D., Beniest, A., Grall, C., Chanier, F., Caroir, F., Dessa, J.-X., Oregioni, D., and Nercessian, A. Evolution of the North Anatolian Fault from a diffuse to a localized shear zone in the North Aegean Sea during the Plio-Pleistocene. *Geophysical Journal International*, 235(3): 2614–2639, Sept. 2023. doi: 10.1093/gji/ggad364.

Sakellariou, D. and Tsampouraki-Kraounaki, K. *Plio-Quaternary Extension and Strike-Slip Tectonics in the Aegean*, page 339–374. Elsevier, 2019. doi: 10.1016/b978-0-12-812064-4.00014-1.

Sakellariou, D., Rousakis, G., Vougioukalakis, G., Ioakim, C., Panagiotopoulos, I., Morfis, I., Zimianitis, E., Athanasoulis, K., Tsampouraki-Kraounaki, K., and Mpardis, D. DEFORMATION PATTERN IN THE WESTERN NORTH AEGEAN TROUGH: PRELIMINARY RESULTS. *Bulletin of the Geological Society of Greece*, 50 (1):124, July 2017. doi: 10.12681/bgsg.11708.

Thingbaijam, K. K. S., Martin Mai, P., and Goda, K. New empirical earthquake source-scaling laws. *Bulletin of the Seismological Society of America*, 107(5):2225–2246, Sept. 2017. doi: 10.1785/0120170017.

Utsu, T. A Statistical Study on the Occurrence of Aftershocks. *The Geophysical Magazine*, 30:521–605, 1961.

Wadati, K. On the Travel Time of Earthquake Waves. *Geophysical Magazine*, 7:101–111, 1933.

Waldhauser, F. *hypoDD-A Program to Compute Double-Difference Hypocenter Locations*. 2001. doi: 10.3133/ofr01113.

Waldhauser, F. and Ellsworth, W. A Double-Difference Earthquake Location Algorithm: Method and Application to the Northern Hayward Fault, California. *Bulletin of the Seismological Society of America*, 90(6):1353–1368, Dec. 2000. doi: 10.1785/0120000006.

Wang, R. A simple orthonormalization method for stable and efficient computation of Green's functions. *Bulletin of the Seismological Society of America*, 89(3):733–741, June 1999. doi: 10.1785/bssa0890030733.

Wells, D. L. and Coppersmith, K. J. New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bulletin of the Seismological Society of America*, 84(4):974–1002, Aug. 1994. doi: 10.1785/bssa0840040974.

Wessel, P., Smith, W. H., Scharroo, R., Luis, J., and Wobbe, F. Generic mapping tools: improved version released. *Eos, Transactions American Geophysical Union*, 94(45):409–410, 2013.

Wiemer, S. and Wyss, M. Minimum Magnitude of Completeness in Earthquake Catalogs: Examples from Alaska, the Western United States, and Japan. *Bulletin of the Seismological Society of America*, 90(4):859–869, Aug. 2000. doi: 10.1785/0119990114.

doi:10.26443/seismica.v3i1.1144

SEISMICA

What does my technology facilitate? A toolbox to help researchers understand the societal impact of emerging technologies in the context of disasters

Lorena Daphna Kuratle *[1,2], **Irina Dallo** [2], **Michèle Marti** [2], **Michael Stauffacher** [1]

[1]ETH Zurich, Transdisciplinarity Lab, Switzerland., [2]ETH Zurich, Swiss Seismological Service, Switzerland.

Author contributions: *Conceptualization*: L.D. Kuratle, I. Dallo, M. Marti, M. Stauffacher. *Data Curation*: L.D. Kuratle. *Formal Analysis*: L.D. Kuratle. *Investigation*: L.D. Kuratle. *Methodology*: L.D. Kuratle, I. Dallo, M. Marti, M. Stauffacher. *Project Administration*: L.D. Kuratle. *Supervision*: I. Dallo. *Validation*: L.D. Kuratle, I. Dallo, M. Marti, M. Stauffacher. *Visualization*: L.D. Kuratle, I. Dallo. *Writing – original draft*: L.D. Kuratle. *Writing – review & editing*: L.D. Kuratle, I. Dallo, M. Marti, M. Stauffacher.

Abstract Disaster risk is increasing globally. Emerging technologies – Artificial Intelligence, Internet of Things, and remote sensing – are becoming more important in supporting disaster risk reduction and enhancing safety culture. Despite their presumed benefits, most research focuses on their technological potential, whereas societal issues are rarely reflected. Taking a societal perspective is vital to ensure that these technologies are developed and operated in ways that benefit societies' resilience, comply with ethical standards, are inclusive, and address potential risks and challenges. Therefore, we were particularly interested in understanding how societal impacts can be considered and leveraged throughout the development process. Based on an explorative literature review, we developed a toolbox for professionals working on emerging technologies in disaster risk reduction. By applying a Delphi study with experts on AI in seismology, we iteratively adapted and tested the toolbox. The results show that there is a need for guided reflection in order to foster discussion on the societal impacts. They further indicate a gap in the common understanding of how a technology is defined and what role it should play in disaster risk reduction. That is crucial for developing inclusive technologies or defining regulations. Our toolbox was found to be useful for professionals in reflecting on their developments and making technologies societally relevant, thereby enhancing societies' resilience.To extend the implementation of the toolbox, it is essential to facilitate additional promotion through avenues such as workshops and conferences. This process should align with the established framework of project management and the policy cycle.

Non-technical summary The frequency and severity of disasters, from both natural hazards such as flash floods and human hazards such as terrorist attacks, are increasing. Newly developed technologies are one way to improve the prevention of and response to these disasters. Recent research has mainly focused on the technological issues of those technologies, with a view to analysing their efficiency. Little research has been conducted to assess whether the technologies help societies in dealing with disasters. This study tries to fill this gap by proposing a toolbox for professionals who work on and with those technologies to help and guide them through a reflection process on what the impact of the technology is on societies. The toolbox was iteratively developed based on a literature review. We tested the toolbox with experts on AI in seismology by using an expert elicitation method (Delphi study). The results show that the toolbox is a helpful starting point for reflection and that the beginning of the discussion needs to be a common understanding on what these technologies are. Only then can the discussion lead to a fruitful further development of the technology to help people deal with disasters.

Production Editor:
Gareth Funning
Handling Editor:
Laure Fallou
Copy & Layout Editor:
Théa Ragon

Received:
October 27, 2023
Accepted:
February 19, 2024
Published:
Mars 1, 2024

1 Introduction

Disaster risk is increasing globally, through both natural and anthropogenic hazards such as earthquakes, wildfires, and terror attacks or chemical accidents (UNDRR, 2022). As the climate crisis evolves, natural hazard events will become more intense and more people will be exposed and negatively affected in the coming decades (IPCC, 2023). Disaster Risk Reduction (DRR) measures are indispensable to mitigate those impacts. The United Nations Office for Disaster Risk Reduction (UNDRR) has formulated the *Sendai Framework for Disaster Risk Reduction* for the period 2015 to 2030 as a response to the need for proper and collaborative actions to address the increasing complexity of disasters (Aitsi-Selmi et al., 2015).

In recent decades, emerging technologies have considerably influenced societies' safety cultures and, consequently, DRR efforts (ITU, 2019). Emerging technologies such as Artificial Intelligence (AI), Internet of Things (IoT), and remote sensing are applied for multiple hazards and for various steps in the disaster management cycle (ITU, 2019), i.e. prevention, planning,

*lorena.kuratle@sed.ethz.ch

and response. Besides enhancing the efficiency and reducing the costs of DRR efforts, emerging technologies can also increase the digital divide, meaning that these technologies are not available for everyone and, thus, could make DRR efforts unjust and only accessible to certain parts of societies (Shaw, 2020).

To date, little research has been conducted on the societal impacts of these technologies for DRR. Lucivero et al. (2011) propose a combination of ethical tools to assess the expectations on these technologies. Gevaert et al. (2021) state that there is a need to explore the societal impacts in order for AI to be fair and just. Gevaert et al. (2021) and Izumi et al. (2019), for example, call for more co-production among researchers and developers when assessing innovation for DRR. This again should not happen during the last mile, but in the first mile of the technology development (Shaw, 2020). Professionals' perspectives and user needs should thus be included from the beginning to enhance DRR efforts.

With our study, we address this research gap by developing a toolbox based on an explorative literature review. The toolbox aims to support professionals (researchers and developers) to reflect on their technologies with regard to their potential societal impacts, answering the question: "What is the potential role of my emerging technology in enhancing safety culture and DRR efforts from a societal perspective?" The toolbox is a set of guiding questions covering the functionality, usability, and societal issues of a technology and can help to identify potential gaps or further need for reflection. We also performed a proof of concept on the example of *AI in seismology* by conducting two rounds of a Delphi study with experts in the field to evaluate the accuracy and usability of the toolbox and answer the following two research questions:

1) Can we iteratively derive a toolbox from literature to support professionals in reflecting on the societal impacts of a technology in order to enhance safety culture within DRR?

2) Does this toolbox support professionals in reflecting on the societal impact of a technology for safety culture within DRR, i) in general and ii) for the example of AI in seismology?

2 State of the Art

2.1 Safety culture and DRR

Disaster risk reduction describes efforts of preventing new and reducing as well as managing already existing risks to enforce resilience (U.N.D.R.R., 2024). Safety culture as part of DRR considers contextual factors and describes "the behaviors and actions of individuals inclusive of decision-makers both public and private, and civil society that reflect a commitment to and are concerned with minimizing risk, injury and losses to human life and the environment" (Marshall, 2020, p5). Safety culture thus describes societal dynamics that are manifested and reproduced in individuals' actions when it comes to safety and encompasses how people deal with disaster and disaster risk and whether they apply safety measures. Consequently, a system, community, or society, which is exposed to any risks and hazards reacts differently depending on its existing safety culture. Therefore, it is crucial to understand local safety culture to enhance DRR and to successfully implement a technology for DRR. If local safety culture is neglected, the implementation of DRR measures may not be successful.

2.2 The role of emerging technologies for DRR and safety culture

In order for technologies to be successfully used, inclusive, and societally relevant, it is crucial to understand safety culture and the influence the technology has on disaster risk reduction. One approach to enhancing DRR and safety culture lies in innovations (Izumi et al., 2019), of which emerging technologies are a part, alongside social innovations such as participation. Emerging technologies for DRR are understood as technologies that are broadly used and have the potential to essentially influence the way societies deal with disasters (Shaw, 2020) and to enhance their resilience (Sakurai and Shaw, 2021). The focus in this study is on AI, IoT, and remote sensing, because these three technologies can be understood as umbrella terms for a broad range of technologies, and, combined, can increase the impact for DRR (e.g. IoT can be combined with AI for predicting hazardous events, Furquim et al., 2018). **AI** refers to Artificial Intelligence and its broad spectrum of applications, e.g. machine learning, deep learning, and natural language processing. **IoT** describes wireless sensor networks that collect data. **Remote sensing** relates to the technology used to study objects from afar, for example with satellites.

2.2.1 The current application of emerging technologies for DRR

In Table 1, we summarize different applications of emerging technologies in DRR, distinguishing between the *technologies* AI, IoT, and remote sensing and the following hazards: terror attacks, flash floods, wildfires, and earthquakes.

Overall, the application and thus the potential of the different technologies for the different hazards do not differ significantly. All of the technologies have the potential to enhance data analysis and processing and, to some extent, forecasting of hazardous events, and are applied before, during, and after disasters. For all three technologies, we found only a few literature studies that assessed the societal impact.

2.2.2 The benefits of emerging technologies

Technological advancements, such as the implementation and development of AI, through for example machine learning or deep learning applications, have created new possibilities for DRR (ITU, 2019). According to ITU (2019), **AI** can improve disaster management by enhancing the recovery and response time. Further, AI is used for different hazards (Surya, 2020; Datta et al., 2022; Khan et al., 2018) and can make disaster management more efficient through faster data analysis, for

	Emerging technologies		
Hazards	Artificial Intelligence	Internet of Things	Remote Sensing
Terror attacks	Detecting potential terror attacks, preventing and predicting mass shootings (Rieland, 2018; Singer, 2022)	Distribution of cheap sensors could help enhance prediction and detection of potential terror attacks, and response to terror attacks in schools or crowds (Gao, 2016; Alsalat et al., 2018).	Counter-terror operations and monitoring applications, e.g. for military operations (Majumdar et al.)
Flash floods	Prediction (Mitra et al., 2016) and forecasting (Costache and Tien Bui, 2019), creating maps for better risk management of flood hazard (Arabameri et al., 2020)	Prediction and monitoring (Furquim et al., 2018; Arshad et al., 2019), management (Goyal et al., 2021)	In combination with machine learning it is used for prediction (Hussein, 2019) and the monitoring and managing of flash floods (Mishra, 2021).
Wildfires	Improvement of early warning and prediction of fire patterns and help with evacuation patterns (Zhao et al., 2020) Prediction of wildfires (Guerrero, 2022).	In combination with remote sensing and machine learning it can improve monitoring (Kaur and Sood, 2019), early detection and warning (Bushnaq et al., 2021; Verma et al., 2021).	Creation of warning maps (Cao et al., 2017) and in combination with machine learning can help to predict fire spread (Huot et al., 2022)
Earthquakes	Improvement of aftershock forecasts and earthquake early warning (Wu et al., 2021)	The use of mobile phones (Zambrano et al., 2017; Wu et al., 2021) can help monitor earthquakes (Taale et al., 2021), and in combination with machine learning improve earthquake early warning.	Analysis of damage after an event (Dong, 2013) and assessment of ground movements (Rathje and Franke, 2016)

Table 1 Summary of the different applications of AI, IoT, and remote sensing for terror attacks, flash floods, earthquakes, and wildfires

example (Sun et al., 2020). AI and big data are also applied to prevent or react to mass shootings and terror attacks (e.g. Staniforth and Akhgar, 2015; Rieland, 2018; Singer, 2022; Ionescu et al., 2020) or to model and predict flash floods (e.g. Costache and Tien Bui, 2019; Arabameri et al., 2020). Further, Mousavi and Beroza (2022) have shown that machine learning, deep learning, and AI applications are already broadly used in seismology and have the potential to significantly influence the field: i) using deep learning for earthquake early warning (EEW; Wu et al., 2021); ii) detecting seismic signals and forecasting seismic activities with machine learning (Seydoux et al., 2020); and, to a lesser extent and even controversially, iii) helping to predict earthquakes (e.g. Banna et al., 2020; Marhain et al., 2021). AI applications are also applied for wild fire predictions and modelling, and evacuation procedures (Zhao et al., 2020). In short, there are promising AI applications in all stages of the disaster cycle (before, during, and after an event) and for multiple hazards.

In order to function properly, AI applications need data (Sun et al., 2020). One popular and cheap way to gather data is to use wireless sensor networks, also called **IoT**. As (Adeel et al., 2018) and (Ray et al., 2017) have shown, IoT is a relevant enabler to enhance disaster monitoring and management for multiple hazards such as earthquakes, terror attacks, and flash floods, often in combination with machine learning or AI applications (e.g. Kaur and Sood, 2019; Goyal et al., 2021). The application of IoT seems very promising for DDR as it can be used in real time, e.g. for early warning and rescue operations (Ray et al., 2017). According to ITU (2019), IoT is also suitable for disaster management because sensors can be applied in multiple settings and for different hazards: they can measure and send signals and warnings from a diverse set of locations (e.g. from trees, from the ground, in buildings).

Another type of technology used for DRR is **remote sensing**. Remote sensing is a technology that is used to study objects from afar (Kaku, 2019), e.g. using satellite data to gather data and information about an area. Remote sensing is particularly helpful for DRR because the acquisition of data can happen very fast and cost-effectively, and cover a large area (e.g. Bello and Aina, 2014; Novellino et al., 2018). This leads to a more effective assessment of an area, both before and after a disaster. Remote sensing can also be applied for different hazards (Bello and Aina, 2014). Mishra (2021), for example, has identified the benefits of real-time monitoring of flash floods through remote sensing and the importance of cheap monitoring possibilities, while Dong (2013) has highlighted the use of remote sensing to evaluate the damage after an earthquake.

2.2.3 The barriers to emerging technologies

Despite the numerous benefits described above, literature indicates possible pitfalls for the use of emerging technologies for DRR (e.g. Bello and Aina, 2014; Sun et al., 2020). Generally, there is a lack of accessibility and integration of ethical and social issues (e.g. Boyd and Crawford, 2012; Crawford and Finn, 2014; Sun et al., 2020). Further, as ITU (2019) describes, there is a lack of standardization and systemization to ensure their broad applicability.

The use of emerging technology can also broaden the digital divide (Shaw, 2020). The digital divide, a term firstly used by Katz and Aspden (1997) describes the phenomenon that technological benefits are not accessible to all but only to certain societal groups (Steyaert and Gould, 2009). To make DRR more inclusive, the digital divide needs to be narrowed (Shaw, 2020). Regarding the challenges of **AI** specifically, Sun et al. (2020) state that many barriers arise due to data-related issues: there is too little or no access to data and there are security or ethical issues. However, too much available data can lead to high computational power required to process it, for instance. Additionally, sometimes the results are not reproducible or not trustworthy and, hence, not helpful for DRR. (Ogie et al., 2018) further argue that a lot of factors, such as local cultures and decision responsibilities must be considered, what needs resources. (Gevaert et al., 2021) also describe how AI in DRR still has unintended ethical issues, e.g. biases arising from the disconnect between the developers and the communities.

As regards **IoT**, it still lacks cost effectiveness, standardization, and context awareness, meaning that in order to harness the full potential of IoT, there is a need for contextual knowledge as well as technological improvements (Ray et al., 2017). Additionally, these systems must become more efficient both in data use and resource management in order to actually enhance disaster management (Adeel et al., 2018).

Remote sensing seems to be especially promising for enhancing disaster preparedness efforts. However, according to Bello and Aina (2014), one major barrier is creating a system that can be applied to different natural and anthropogenic hazards. Additionally, the timely provision of data proves to be challenging (Bello and Aina, 2014). Novellino et al. (2018) have shown that remote sensing is already applied broadly, with the main challenge being field verification (i.e. the inclusion of people affected).

2.3 The societal issues of emerging technologies

As mentioned above, societal issues have so far been broadly neglected in the assessment of emerging technologies' potential for DRR. This is also confirmed by a review study on universal design, referring to designs that are usable by everyone with a maximal benefit (Connell et al., 1997). Gjøsæter et al. (2020) conclude that despite the efforts of making ICT emergency technologies more accessible, there is still a gap to design those technologies for everyone, i.e. every possible user. Additionally, they highlight that the needs of diverse stakeholders and a human-centered approach should be included in the design of technologies for emergency management. For instance, Petersen et al. (2023) and Dallo et al. (2022) chose such a path in their research by including relevant stakeholders in the design of hazard and risk communication products. This approach of co-production allows to enhance usability between developers and users to ensure user-centred communication, which is necessary for effective hazard and risk communication as well as the usability of a technology to move from a last-mile to a first-mile approach (Shaw, 2020). Some scholars such as Petersen et al. (2023) argue to include the ethical, legal and social issues in the assessment of those technologies in a nuanced way in order to actually address them. With respect to usability, end users need to have a positive perception of and trust in a technology in order to apply it and accept the decisions derived from the outputs of these technologies (Kankanamge et al., 2021). Additionally, the technologies need to fit into existing structures such as established communication networks, and the local safety culture, and reflect people's capacities and needs. While there are studies about public perception of emerging technologies in general (e.g. on AI: Kelley et al., 2021), there is little literature on the public perception of their use for DRR. The acceptance and support thereof have thus not yet been elicited.

Another societal aspect is inclusiveness, i.e. consideration of the inclusion of vulnerable groups. One way to be more inclusive is to adopt an intersectional approach (Crenshaw, 1991; Vickery, 2017). Applied to DRR, the intersectional approach helps to find the most vulnerable and marginalized groups (people of colour, immigrants, sick, old, disables, queer etc. people) in different disaster contexts by assessing and uncovering intersecting traits or social variables. (Vickery, 2017). The homogenized term "vulnerable" can lead to a neglect of characteristics and traits that have an influence on the outcomes of a disaster response (Vickery, 2017). It is important to acknowledge that every person can be made vulnerable in a disaster, and that this is contextual. Thus, also the International Organization for Standardization (ISO) includes the personal circumstances in the assessment of vulnerability (ISO 22395, 2018). Intersectional awareness helps to understand vulnerability better.

With our study, we aim to close the still existing research gap of including users and considering ethical implications, by providing professionals (researchers and developers of these technologies) with guidance for thinking about the impact of their technology on societies and for the contextual safety culture. To this end, we focused on two specific societal issues: (i) the user-centred perspective in terms of the usability of a technology; and (ii) inclusiveness, i.e. who benefits from a technology and who is excluded.

3 Methods and Material

The methods used for this study are shown in Figure 1. Based on an extensive literature review (section 3.1), we iteratively developed a toolbox addressing the relevant issues when evaluating the potential of emerging technologies for DRR. Afterwards, we conducted a proof of concept by applying a Delphi study with two survey rounds, which allowed us to improve the toolbox based on expert feedback (section 3.2).

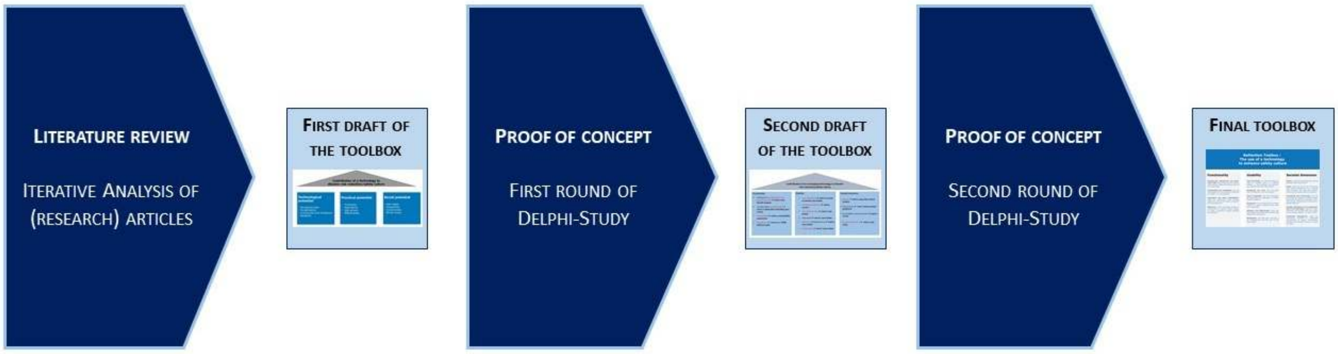


Figure 1 Overview of the methodological procedure: literature review, development and adjustment of the toolbox, and proof of concept with a Delphi study.

	Emerging technologies		
Hazards	**Artificial Intelligence**	**Internet of Things**	**Remote Sensing**
Terror attacks	6	5	2
Flash floods	4	4	2
Wildfires	4	3	2
Earthquakes	4	4	3

Table 2 Number of articles analysed for the four hazards and three technologies [in total 43 articles]

Overarching categories		
Technological potential	**Practical potential**	**Social potential**
Development costs	Practicality	User needs
Transferability	Applicability	Accessibility
Functionality	User groups	Inclusiveness
Reliability	Effectiveness	Ethical issues

Table 3 The three overarching categories and their associated sub-topics

3.1 Literature review for iterative toolbox development

We conducted an explorative literature review based on a search with a number of hazard keywords – earthquake(s), flash flood(s), wildfire(s), terror attack(s), disaster(s) – in combination with disaster risk reduction or disaster management or safety culture or emerging technologies. We searched on the platforms Google Scholar and Web of Science, and applied a "snowballing" method, i.e. looking at the references of the identified literature to access more relevant studies (Greenhalgh and Peacock, 2005). With this literature review, we mapped the current state of the art for the role of emerging technologies in DRR. Based on this, we then iteratively and deductively developed our first toolbox draft. It should be noted that the literature review yielded only a small number of publications overall (see Table 2). In order to gain a broad overview, we first searched for general literature on technologies used for DRR and specifically on the societal impact of those technologies. We found that there was a clear tendency towards the assessment of functionality and sometimes users, but little literature on the societal impact, which is why this became one focus within our toolbox. In order to holistically grasp the potential of an emerging technology to enhance safety culture, we focused not only on the societal issues but also on the practical and technological issues, since these are very interdependent.

Based on the findings and insights from the literature review, we developed a first version of the toolbox. The derived relevant issues for assessing the potential role of emerging technologies to enhance safety culture were organized in three overarching categories –technological, practical, and social potential – each with four associated sub-topics.

In a second round, the chosen studies were analysed to understand what findings, if any, each study provided with respect to these categories. In every article, we examined whether or not each of the categories was assessed.

This iterative procedure combined with discussions with fellow researchers allowed us to complement aspects and to merge certain issues. This led to a first draft of the toolbox ready to be tested in a proof-of-

	Delphi study	
Socio-demographics	Survey – Round 1	Survey – Round 2
# participants	12	7
Average age	37 years	39 years
Gender	n =8: male n =2: female n =2: do not wish to disclose	n =5: male n =2: female
Place of work	n =1: prefer not to say n =1: China n =2: United Kingdom n =2: USA n =2: France n =4: Switzerland	n =1: France n =3: United Kingdom n =3: Switzerland
Years in current position	n =1: 5-10 years n =2: 10-20 years n =3: more than 20 years n =6: 1-5 years	n =2: 10-20 years n =5: 1-5 years
Level of expertise	n =1: no expertise n =1: high expertise n =2: very low expertise n =8: medium expertise	n =1: no expertise n =1: high expertise n =5: medium expertise
Research focus	n =1: earthquake forecasting n =1: earthquake prediction n =2: none n =4: earthquake early warning n =4: rapid impact assessment	n =1: earthquake prediction n =1 rapid impact assessment n =4 earthquake early warning

Table 4 Characteristics of participants in the first and second survey rounds

concept study to determine whether it actually allowed professionals to reflect on the potential of an emerging technology for DRR. For this we chose the Delphi study method.

3.2 Delphi study to test the toolbox

By means of a Delphi study, we conducted a proof of concept of our toolbox and assessed the potential of AI in seismology. Experts on AI in seismology were recruited based on their proven expertise in the field and invited to participate in two survey rounds using the online survey tool *Unipark*(more information about the recruitment and participants can be found in the next section). The tool allows for simple, location-independent, anonymous participation. Anonymity of the participants is one key characteristic and advantage of the Delphi study because it reduces the risks of individuals dominating group discussions, thus pre-empting manipulation and coercion (Dalkey, 1972).

In both rounds, participants had to rate different statements (see Sections 3.2.2 and 3.2.3) on a 5-point Likert scale, from 1=strongly disagree to 5=strongly agree. We also included open-ended questions to let them comment on their ratings. Based on the comments provided in the first round, we adjusted or added new statements to be rated in the second round (Table 3).

3.2.1 Participants

For the expert recruitment, we chose to invite around 90 participants via email. Our selection criterion was that the possible candidates must have written a peer-reviewed article on AI in seismology within the last three years. The target was to reach about 15-30 experts, since this is the number recruited in most Delphi-studies (Hsu and Sandford, 2007). Further, we aimed to reach involve a diverse group of experts. We tried to counterbalance the Eurocentric bias by inviting experts from all over the world, specifically also targeting scientists based on the Asian and African continents, and by inviting experts of all genders. The first attempt to recruit experts via email was not sufficiently successful. We thus contacted experts through our project networks and the expert pool of the Swiss Seismological Service at ETH Zurich. In the end, we had a total of 12 experts who completed the first survey and 7 experts completing the second survey. The socio-demographics of the participants are summarized in Table 4.

3.2.2 First survey round

In the first round, we asked the participants to rank 66 statements about the toolbox and its applicability in assessing the potential of AI in seismology. The statements for the specific case were derived from the literature and discussions with seismologists. The statements tried to encapsulate the state of the art of AI in seismology, which can be summarized as follows:

AI in seismology is used for *fast data processing* (Mousavi and Beroza, 2023). This is especially promising because it seems to be cheaper than previous procedures used for modelling (Essam et al., 2021). AI can help enhance *EEW* (Meier et al., 2020; Iaccarino et al., 2021; Wu et al., 2021; Datta et al., 2022) and improve methods to *forecast earthquakes* (Mousavi and Beroza, 2023; Beroza et al., 2021) using e.g. deep learning (Seydoux et al., 2020). AI is also used for *rapid impact assessments* (Harirchian et al., 2021; Stojadinović et al., 2021). Some scholars even argue that AI can be used to predict earthquakes (e.g., Marhain et al., 2021), but this is heavily disputed since predicting the precise location, time, and magnitude of a future earthquake is not possible at the current state.

Thus, AI in seismology, with its manifold potential in DRR, which is still at an early stage of implementation, offers an ideal case study. Further, the gap in the elicitation of the societal impact is also problematic in this domain.

The rating of the statements was followed by open-ended questions addressing the experts' understanding of AI in seismology and their general opinion on the pillars of the toolbox. We also assessed the experts' age, gender, job, and location of research in order to ensure a diverse set of participants. The survey was pre-tested by a seismologist and several experts in social sciences.

For the data analysis, we followed the procedure of Vogel et al. (2019) and used SPSS. While the socio-demographics were analysed descriptively, the statements were analysed using percentages. We assumed that consensus about a statement was reached when more than 70% of the participants gave answers according to the categories defined as agreement and disagreement (Vogel et al., 2019). We defined these by adding categories 1 and 2 ("Don't agree at all", "Don't agree") to indicate disagreement, and categories 4 and 5 ("Agree" and "Fully agree") to indicate agreement. Category 3 indicated a neutral position. The open-ended questions were analysed qualitatively using Word and an inductive approach. After analysing the first survey, both quantitatively and qualitatively, we found that there was little to no consensus on the statements, which made it difficult to adapt them. Therefore, we chose to adapt the toolbox based on the insights from the qualitative analysis and exclude the statements as a whole from the second survey. This is consistent with the Delphi study procedure, as defined by Pohl (2020), since we followed an iterative process and adapted the survey after each round based on the experts' answers.

3.2.3 Second survey round

In the second round, the experts from the first round were asked to evaluate the adapted toolbox and to comment on a concise definition of AI in seismology. The survey consisted of three parts: (i) a shared definition of AI; (ii) the adapted toolbox; and (iii) demographic information. For the data analysis, we again followed the procedure of Vogel et al. (2019).

4 Results

In sections 4.1 to 4.3, we describe the results of the explorative literature review and iterative toolbox development (section 4.1) and the first (section 4.2) and second (section 4.3) rounds of the Delphi study survey. The in-depth literature review can be found in the supplement in the tables lr1, lr2, and lr 3. The results of the Delphi study follow the structure of the surveys, starting with the findings for the toolbox in general and then the specific case *AI in seismology* (Supplement, Delphi Survey (DS) – Round 1). These results show the in-depth answer to the research questions (section 1) concerning whether the developed toolbox is applicable and suitable for professionals to reflect upon the potential role of an emerging technology to enhance safety culture.

4.1 Results of the explorative literature review and iterative toolbox development

With our literature review, we iteratively and inductively developed the first solid draft of our toolbox to test in a Delphi study (Figure 2). The first draft consisted of three distinct pillars with four categories in each, all of which can be assessed individually, as shown in Figure 2. The goal was to holistically cover the role of an emerging technology in enhancing safety culture. In addition to the technological and practical issues, we aimed to elicit the societal impact of a technology, because very little literature was found on this.

4.1.1 Definition of AI

In the first part of the survey, we asked the experts to define AI in general and to explain what they thought was the potential of AI for DRR. We identified three common thoughts: i) AI is a term used to describe computational processes that involve learning; ii) AI mimics human intelligence; and iii) AI is able to process information fast. Nine out of 12 experts also agreed that AI could help enhance DRR, but that this potential should be assessed when AI has developed further. Based on these findings, we derived a definition for AI in seismology, which we then presented in the second survey round for the experts to comment on.

4.1.2 Feedback on toolbox

In the second part of the survey, we presented the first draft of the toolbox (see Figure 2). Concerning the toolbox as a whole, 6 out of 12 experts found it difficult to understand in which context and for what purpose the toolbox would be used and what the concrete implementation would look like. However, 7 out of 12 experts stated that it was a good starting point with room for improvement when it came to context, objectives, and specific items within the pillars. Additionally, a general feedback was that the metrics for all the categories within each pillar should be added, as the following statement shows: *"I like these categories and I believe they are well described. But it will be hard to quantify how transferable or how limited a technology is* (ID10)."

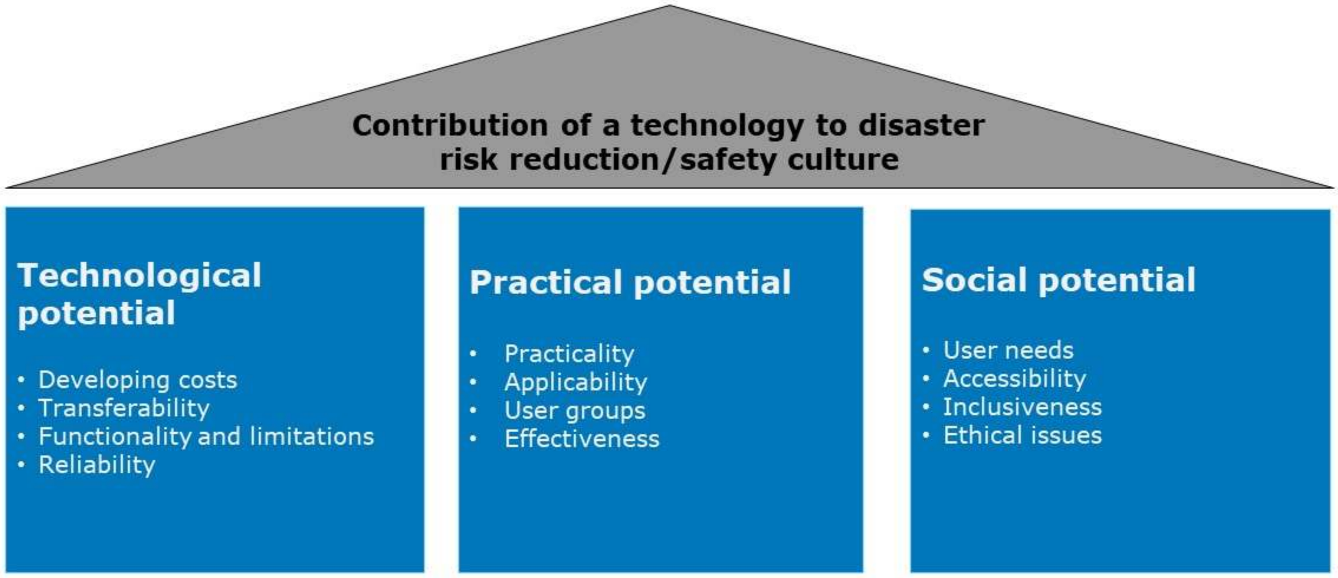


Figure 2 First draft of the toolbox after the literature review for the first survey round of the Delphi study. The toolbox was derived from the literature review. It consists of three pillars – technological potential, practical potential, and social potential. Within these pillars, there are categories that can be assessed independently in order to understand the respective pillar.

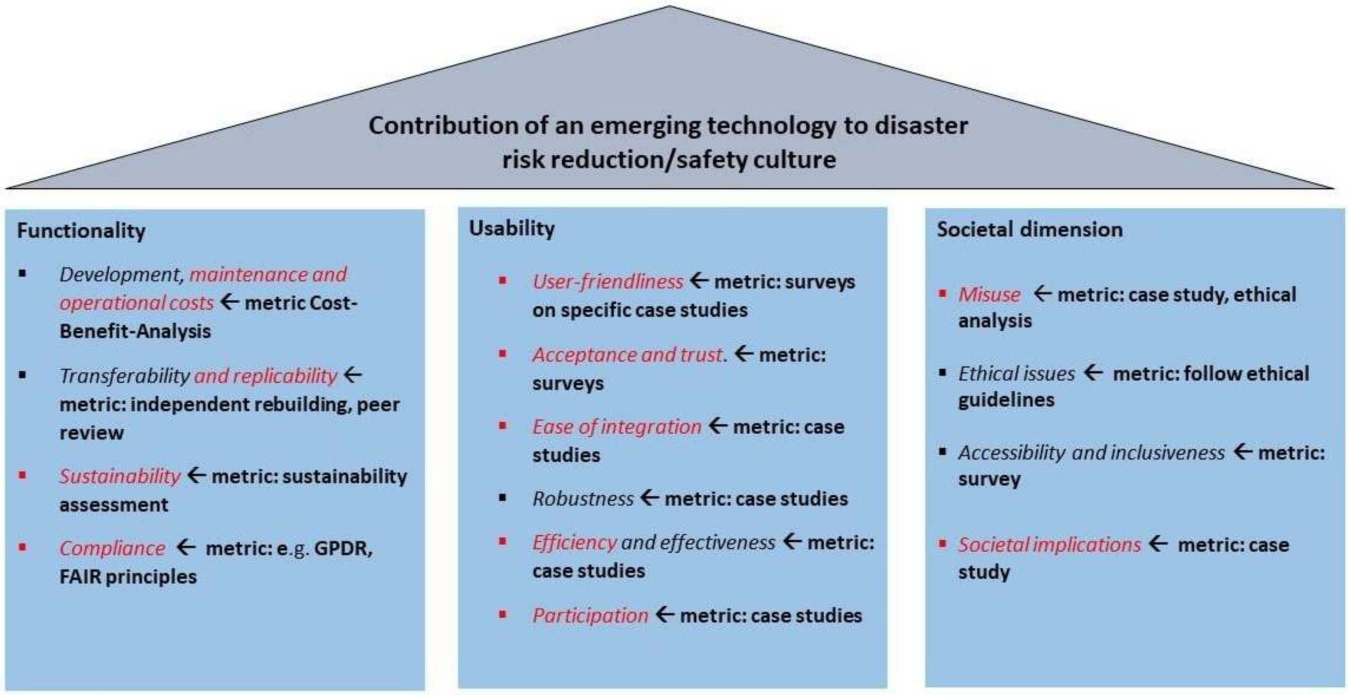


Figure 3 Adapted toolbox based on first round of the Delphi study. Changes are highlighted in red. The adapted toolbox was presented in this form in the second Delphi study survey round.

We then asked the experts to comment on each pillar separately. For each pillar, the experts provided general feedback and were able to suggest additional issues that in their opinion were missing.

For the *technological potential* pillar, in addition to better clarifying the purpose and providing metrics, there was a consensus that the "Developing costs" category should be expanded to include maintenance and operational costs as well as benefits, as the following statement underlines: *"[...]. For example cost has no sense if benefit is not added* (ID2)." Besides these costs, the participants mentioned other factors to be added to the pillar, such as sustainability (see Figure 3).

For the *practical potential* pillar, the feedback was similar. Participants said they would like to have more context in order to better understand the application of the toolbox. They also suggested extending the focus on end users, which was considered in the second draft (Figure 3). The feedback on this pillar also included the question of what role users should play within the development of AI in seismology.

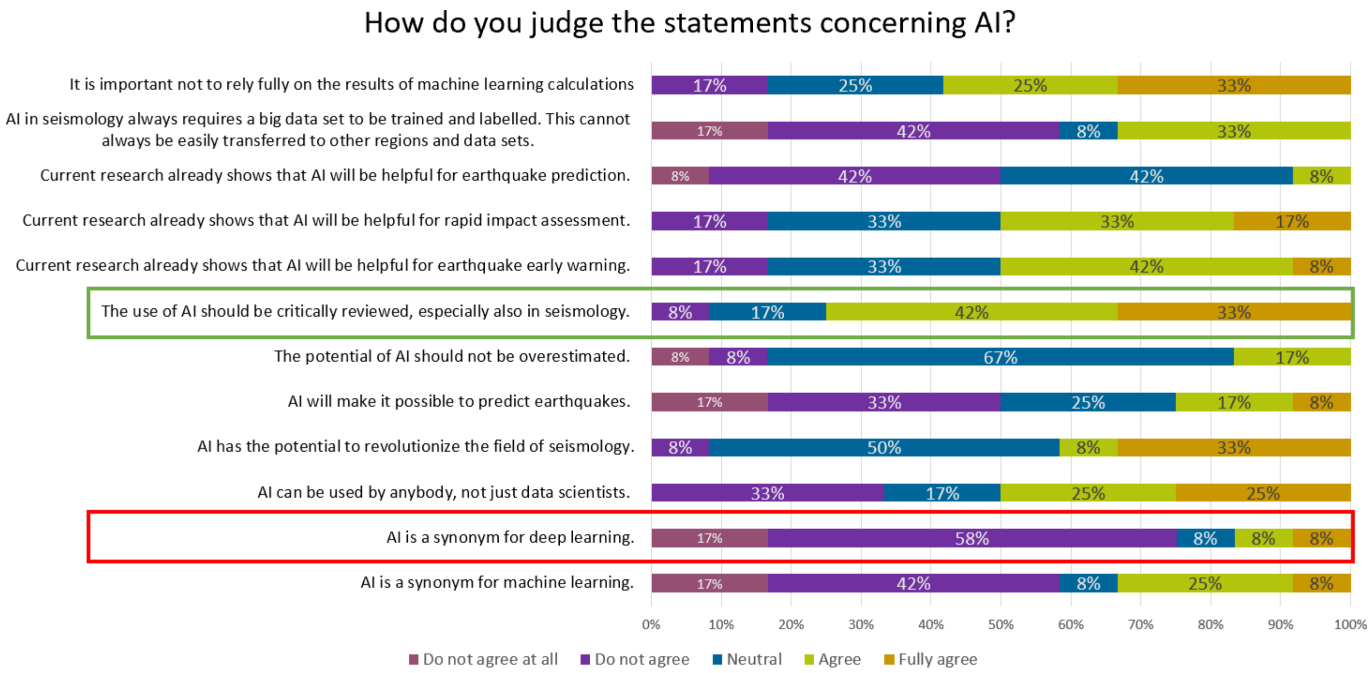


Figure 4 Statements related to AI in general: agreement is highlighted in green and disagreement in red (>70% agree or disagree).

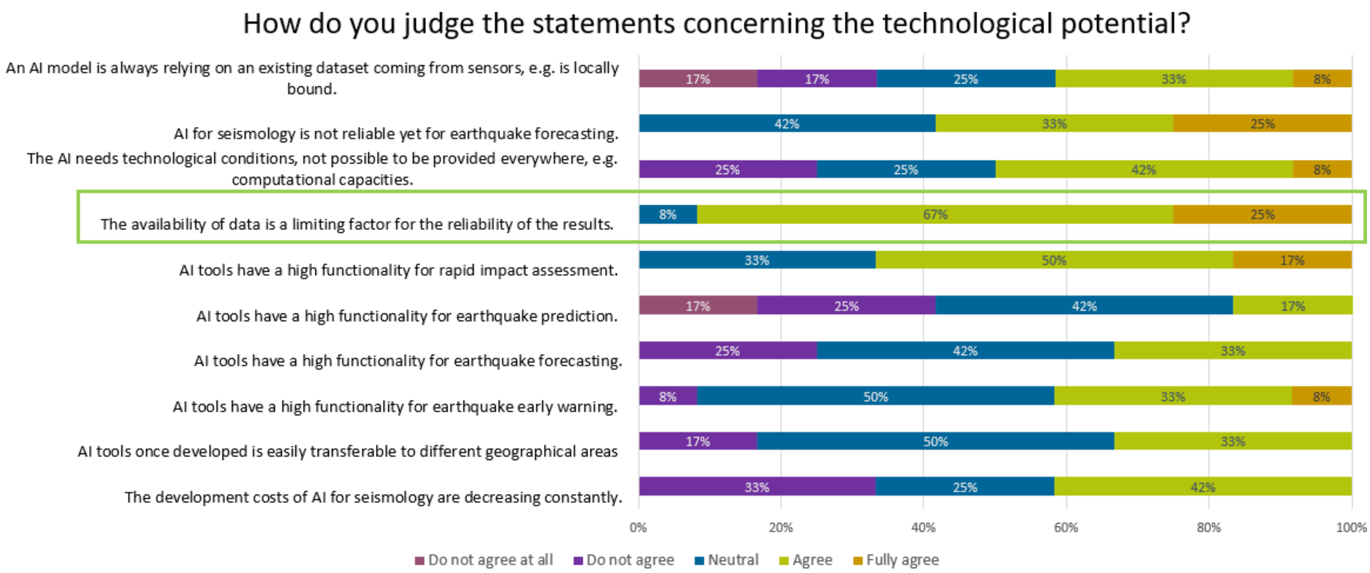


Figure 5 Statements concerning the technological potential: agreement is highlighted in green (>70% agree).

As one participant put it: *"It probably really depends on the intended users/usage scenarios: 'recognize' a disaster from data; 'characterize' the disaster and its potential; support decision making (different stakeholders/users) – I find it impossible to assess practicability 'in general'* (ID9).*"*

4.1.3 Adapted toolbox

Based on the feedback from the experts in the first survey round, we renamed the pillars: technological potential became *functionality*, practical potential became *usability*, and social potential was renamed *societal dimension*. Additionally, we added further relevant categories and metrics to assess them, and a description of the purpose and goals of the toolbox. This helped us strengthen our main goal of providing experts with guiding questions to assess the societal impact of a technology they are developing (see Figure 3).

The comments concerning the *social potential* pillar were very diverse. On the one hand, the need for ethical considerations in the development of these technologies was highlighted. On the other hand, the feedback was that the ethical considerations differ depending on the role of the end users. Some participants also reflected on the responsibility of the different actors, as the following statement shows: *"None of this is related to the tech itself but to the way it is used by the operator. It is unfair to blame the developer of tech for these issues* (ID8).*"* These issues were added to the toolbox, as shown in Figure 3.

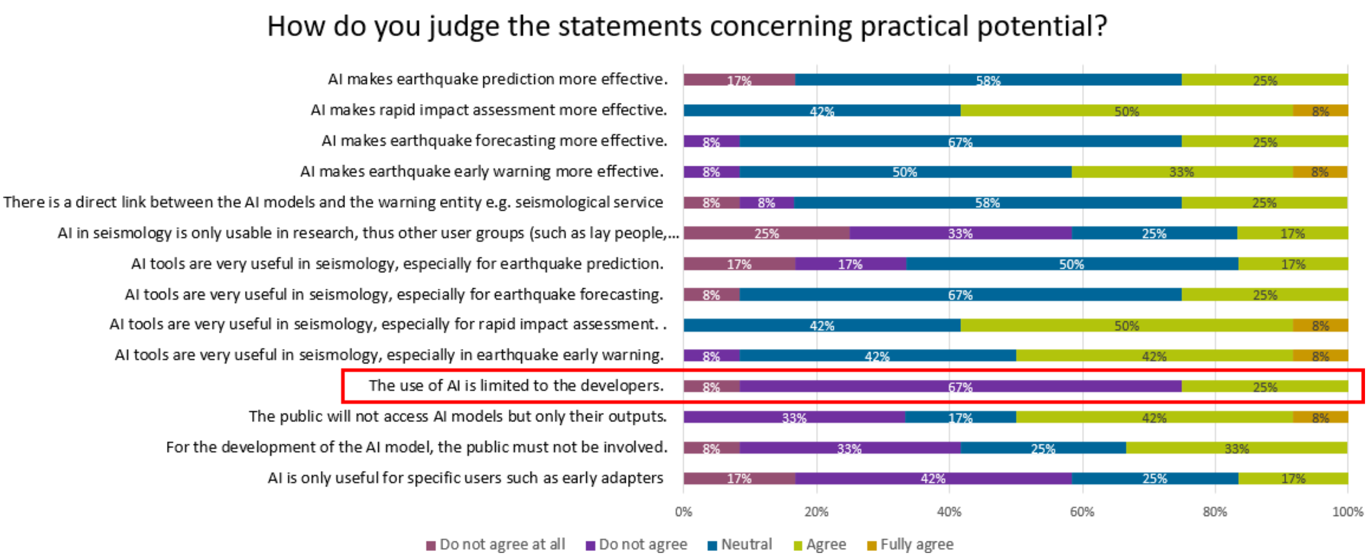


Figure 6 Statements concerning the practical potential: disagreement is highlighted in red (>70% disagree).

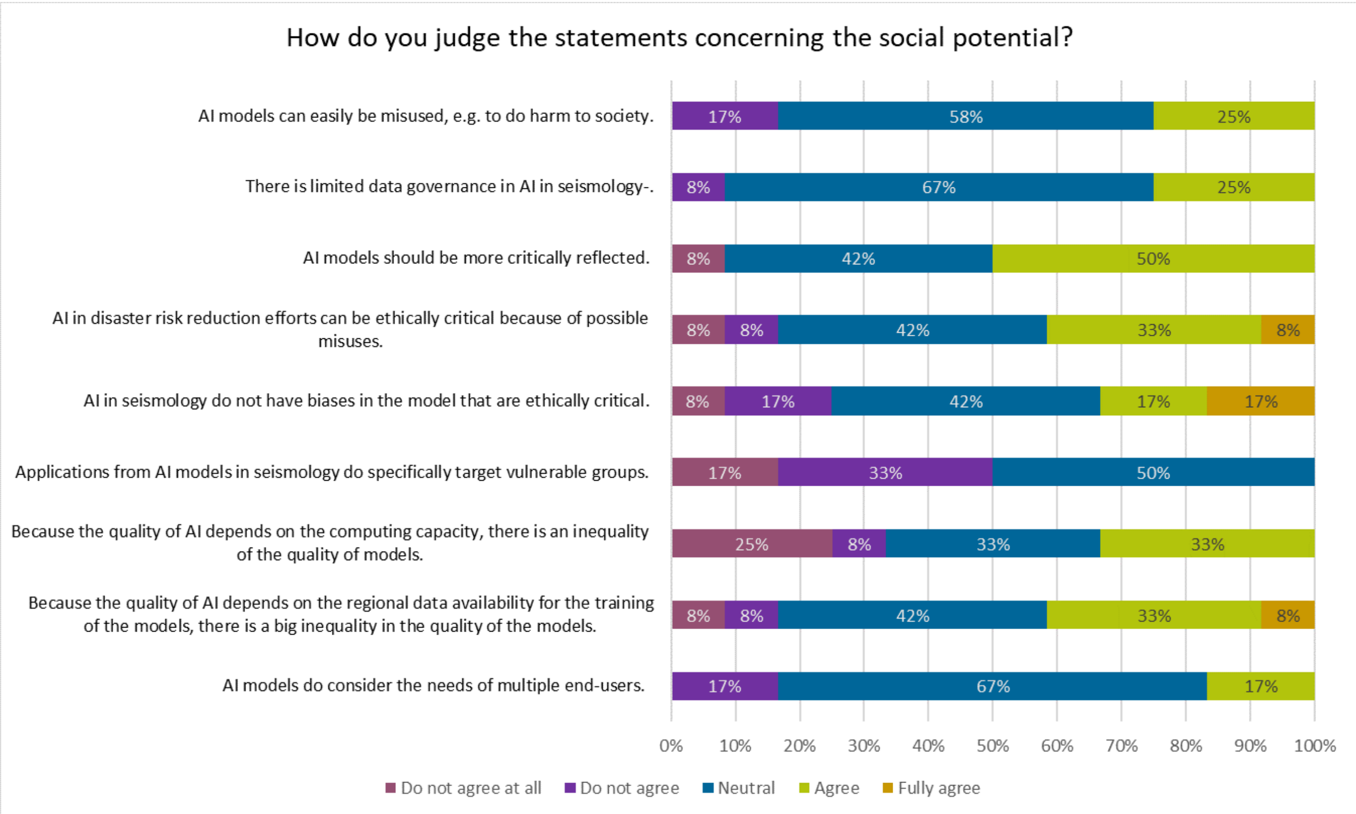


Figure 7 Statements concerning the social potential

4.1.4 The case study

In the third part of the survey, we asked specific questions concerning AI in seismology and wanted to test whether our toolbox actually helped the experts to reflect upon the societal potential of AI in seismology. The question *In which fields of seismology does AI play an important role?* elicited very diverse answers, ranging from communication and earthquake early warning to prediction. However, there was a consensus that the use of AI is still in its beginnings in seismology and that the acceleration of data processing will lead to more achievements. On the question of the greatest potential, the answers ranged from data processing and predicting damage to earthquake prediction.

We also presented statements on the potential of AI in seismology to ascertain whether the experts found the toolbox applicable.

The first batch of statements focused on the general potential of AI in seismology. It is notable that there were only two instances of clear consensus (see Figure 4). The experts agreed on the statement that the use of AI should be critically reviewed, especially in seismology. Further, there was consensus that AI is not a synonym for deep learning. The statement that AI is a synonym for machine learning came close to achieving

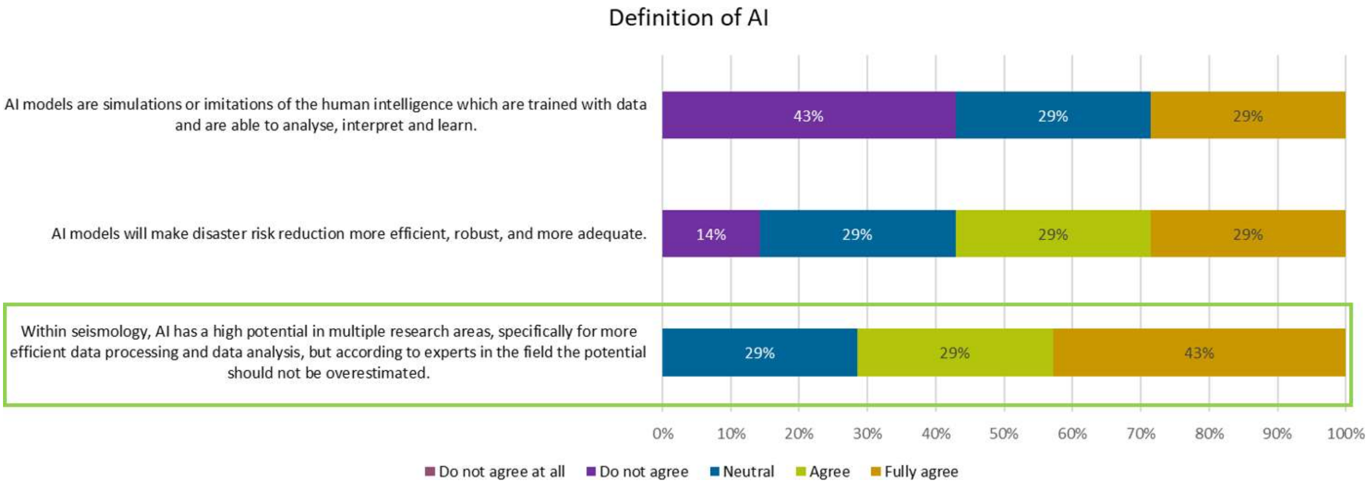


Figure 8 Definition of AI in seismology: consensus is highlighted in green (>70% agree).

a consensus (59%). For the other statements, there were few indications of a consensus, which suggests that the experts do not agree on the potential of AI in seismology and also that the applications of AI in seismology are very broad.

Because there was only very little consensus and a lot of dissent, we chose not to include these statements in the second round. However, the results of this first batch were used to formulate the proposed definition of AI for the second round of the Delphi study.

The statements concerning the three different pillars showed only little consensus.

Concerning the technological potential, there was only one instance of agreement, namely that the availability of data is a limiting factor for the reliability of the results (see Figure 5). The statement "AI has a high functionality for rapid impact assessment" almost reached consensus (67%). Some of the answers show a high degree of neutrality (over 40%).

For the practical potential, the experts agreed that the use of AI is not limited to the developers (Figure 6). No other statement achieved a consensus. The closest to consensus was for the statement "AI is only useful for specific users such as early adapters" (57%). For this user-focused category, there were even more neutral answers than for the technological potential.

For the statements concerning social potential, no consensus was reached. The statements that came closest to a consensus were "AI models should be more critically reflected" and "Applications from AI models in seismology do specifically target vulnerable groups" (50%). All statements attracted over 30% neutral answers (Figure 7).

4.2 Results of the Delphi study – Second round

In the second round, 7 of the initial 12 participants filled out the survey (Table 4). The second survey (Supplement, Delphi Survey – Round 2) was shorter because we chose to focus on the possible common definition of AI in seismology and the adapted toolbox as a whole (Figure 3). Consequently, the experts were no longer asked to rate statements on the specific case of AI in seismology. The new toolbox was again broadly commented on and the changes were viewed positively. One participant called it *"a good start for reflection* (ID7)*"*, which was exactly our goal. Still, several experts suggested some reformulations, renaming, and adjustments, which led to the final toolbox as shown in Figure 9.

4.2.1 Definition of AI

The definition of AI compiled from the answers of the first survey round consisted of three parts, which the experts were able to judge separately (Figure 8):

1) AI models are simulations or imitations of the human intelligence which are trained with data and are able to analyse, interpret and learn;

2) AI models will make disaster risk reduction more efficient, robust and more adequate;

3) Within seismology, AI has a high potential in multiple research areas, specifically for more efficient data processing and data analysis, but according to experts in the field the potential should not be overestimated.

Only for the third part of the definition was a consensus reached. The experts agreed that AI has a high potential within seismology but should not be overestimated.

Multiple comments additionally suggested that the focus of AI in seismology does not lie in the imitation of the human brain but rather in computational implementation and data processing. The following comment by a participant illustrates this: *"I have an issue still with linking AI to 'simulating/imitating human intelligence' - I am not an expert in human intelligence, but I believe that, what exactly that is, is still debated... therefore I would rather describe AI as computational implementations of models of learning/reasoning/concluding that are shaped after current understanding of neural (brain) networks (or so...)* (ID7).*"*

4.3 Final toolbox

After analysing the results of the second round, we made two big changes. Firstly, we chose to solely analyse the enhancement of safety culture and not safety culture within DRR and not DRR overall. The rationale

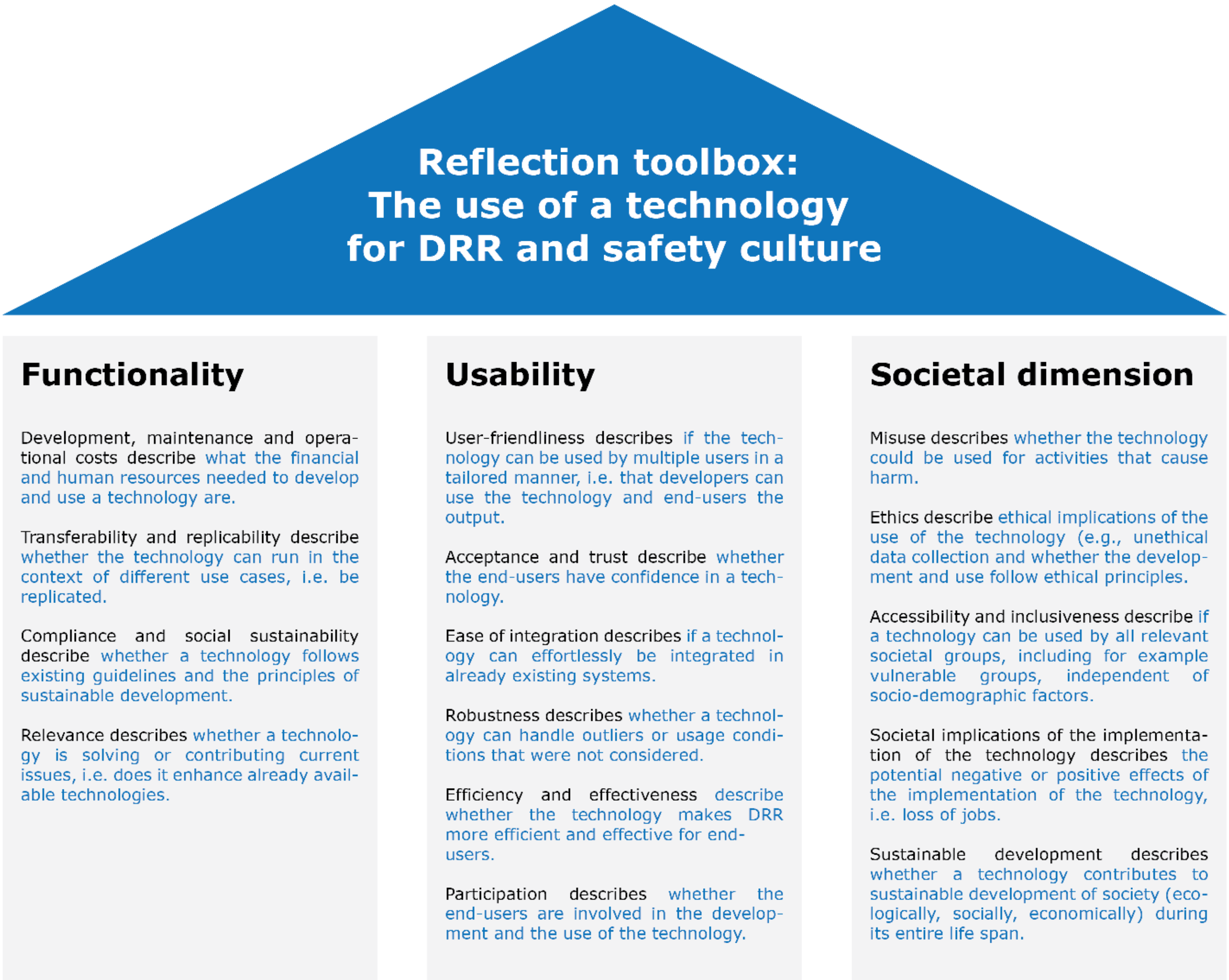


Figure 9 Toolbox to assess the potential of emerging technologies for DRR and safety culture developed in an iterative process consisting of literature review and two Delphi study survey rounds

behind this was that only if the contextual safety culture is improved, DRR effort are/become effective, as realized within the literature review. (see Safety culture and DRR). Secondly, we chose to remove the metrics from the categories. The reasons were that the toolbox should be directly applicable and not require in-depth studies for each category in each pillar. To this end, we formulated questions that can be answered for the analysis (Figure 9). To summarize, we again formulate the main goal of each pillar.

The final toolbox (Figure 9) consists of the following three pillars designed to holistically assess the potential role of an emerging technology in enhancing safety culture:

Functionality: Does it work?
The *functionality* pillar describes whether a technology functions properly during its whole lifespan and contributes to enhancing existing DRR efforts. It can be evaluated by testing the technology in existing applications in laboratory or real-world settings.

Usability: Is it used/usable?
The *usability* pillar describes whether a technology is usable and applicable by different targeted end users (context-independent), and specifically assesses the active use and the intended use of a technology.

Societal dimension: What does it mean for society?
The *societal dimension* pillar analyses the contribution of AI to DRR from a societal and ethical perspective. It addresses possible ethical issues such as misuse of the technology.

5 Discussion

Based on a literature review and a Delphi study, we were able to develop a toolbox to support professionals (developers and researchers) in the systematic reflection on the societal impact of the technology they are developing, implementing, or operating, considering safety culture in order to improve disaster risk reduction.

In the following, we explain how the iterative steps of the Delphi-study has confirmed our findings of the literature review (section 5.1). Further, we discuss how our toolbox could be applied within the project and policy cycle in order to ensure the effective use of the toolbox (section 5.2). Last, we critically reflect on the limitations of our study and discuss future research (section 5.3).

5.1 The comparison of the literature review and Delphi-study

Our toolbox is designed to help professionals to reflect on the technologies' contribution to enhancing societal benefits, encouraging collective actions towards an enhanced safety culture and including marginalized groups within society. The importance of including societal issues emerged from both the literature review and the Delphi study. Past research on the potential of technologies for DRR has mainly focused on the functionality and the usability of those and thereby neglected the societal perspective and their impact on safety culture. The insights from the Delphi study support this finding, with the statements about the technological and practical potential generating most consensus. At the same time, fewer neutral answers were given in these areas (see Figure 6 and Figure 7), indicating a shared scientific understanding.

The International Telecommunication Union (ITU, 2019) conclude in their assessment that disruptive emerging technologies for DRR are improving disaster management but that further research is required to ensure large-scale impacts. With particular regard to increasing societal impacts, they recommend fostering public outreach, i.e. consideration of the purpose and specific target audience, and partnerships between academia and the private sector to improve disaster management overall (ITU, 2019). This is also stressed in the literature review of Gjøsæter et al. (2020) In addition, our study shows that experts are interested in reflecting on their technologies, but emphasize that this is not just their responsibility but the task of all actors involved in the development, implementation, deployment, and use of a technology. This is indicated by the neutral answers for the practical and social potential statements (see Figure 6 and Figure 7). Our toolbox thus consists of questions that are applicable for all actors involved.

The literature review demonstrated that clear definitions of the technologies looked at are lacking: the applications of AI, IoT, and remote sensing are very broad and this is why there is only a tendency towards a common understanding. However, distinct definitions are required in order to be able to discuss the societal impacts of a technology. Consequently, a common understanding needs to be strengthened through further societal and scientific cooperation. This will form the basis for, among other things, drawing up regulations and policies for the development and application of AI (Harasimiuk and Braun, 2021), IoT and remote sensing in order to enhance safety culture.

It is therefore not surprising that AI in seismology is also lacking a common definition, as hinted by the literature review and the Delphi study. Despite the fact that most respondents called themselves experts on AI in seismology, they did not provide the same definition. Given the broad range of possible applications of AI in seismology and the different specializations of the respondents, this seems logical (e.g. Mousavi and Beroza, 2023). Still, the results show that the experts agree on some of the potential and the limitations of AI in seismology. Hence, AI in seismology cannot be reduced to just a single definition but rather should be discussed in the context of each application, with its limitations and pitfalls, and should not be overestimated (Mousavi and Beroza, 2022). In order to understand the potential of AI in seismology to enhance safety culture, the first step should be to understand which specific application of a technology is discussed. Given the variety of definitions, the toolbox and its categories are kept broad, while still serving as a catalyst for critical reflection on the issues under discussion and enabling an assessment of the potential in each specific application.

Still, the comparison of our literature review and the Delphi study shows that we were able to iteratively derive a toolbox which can support professionals in reflecting the societal impacts for safety culture of the technology they use. The specific case study of AI has shown that the toolbox does support professionals.

5.2 The implementation of the toolbox

To reach the purpose of being further developed, the toolbox should be actively used. This can only be achieved if the toolbox is known. One possibility would be organizing workshops with practitioners, by doing more outreach, possibly with the ITU, in order to ensure further development and, in the end, possibly standardization.

Further, existing research indicates that co-production of knowledge is required to improve DRR measures (Ismail-Zadeh et al., 2016; Izumi et al., 2019), i.e. involving stakeholders from the beginning following the first-mile principle (Shaw, 2020) and strengthening the collaboration between science and society (ITU, 2019). The evaluation of the three pillars – functionality, usability, and societal dimension – of our toolbox within the Delphi-study indicates the same: there is a need for a guided discussion and reflection on the consequences of a technology in the scientific community as well as societies to increase awareness, which the toolbox can facilitate by guiding relevant stakeholders in their reflection from the outset.

Once the toolbox is known, potential areas of influence must be identified. To this end, we linked the elements of the toolbox to the policy cycle adapted from Schubert and Klein (2020), as well as the project cycle adapted from the European Commission (2004); see Figure 10.

Setting the agenda firstly is crucial in the project initiation: in this step the goal to enhance safety culture is manifested, and hence the goal to use the toolbox in the process. With the second step, the formulation of the policy, the different foci of the use of the technology and thus the application of the different pillars of the toolbox is chosen. This then leads to the third step, the decision where the time to reflect is spent. In the two final steps, the implementation and the evaluation of the technology happens, once again with the reflection guidance of the toolbox. All these steps happen cooperatively, co-productively, and iteratively, both first-mile to last-mile.

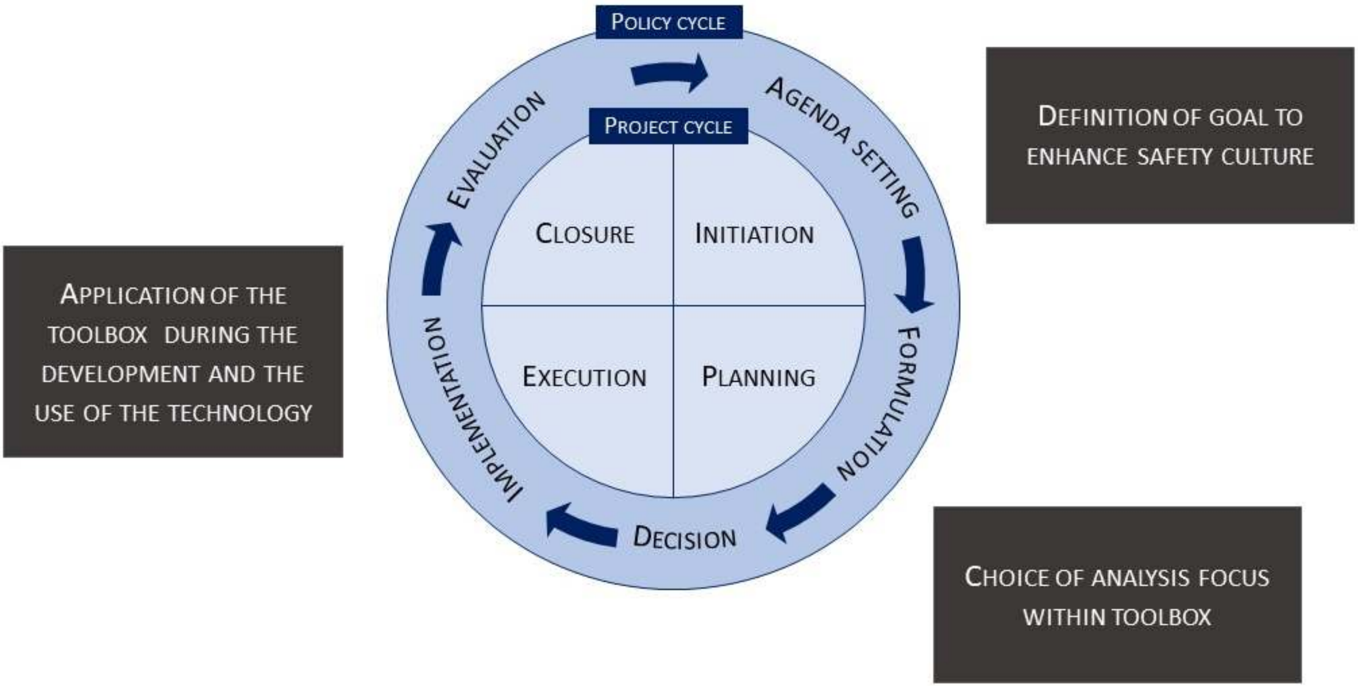


Figure 10 Application of the toolbox (black squares) in the policy cycle (adapted from Schubert and Klein, 2020, blue arrows) and the project management cycle (adapted from European Commission, 2004, blue squares).

5.3 Limitations and next steps

Our study has several limitations that could be addressed in future research.

Our explorative literature review was not conducted fully systematically but rather iteratively, meaning that there was a broad timeframe and limited sample chosen. However, the literature review was solely needed to identify the categories forming the basis of the toolbox and to grasp the state of the art of these technologies in DRR and to then develop the first solid draft of the toolbox. Further, through the expert elicitation (Delphi study), we aimed to overcome these issues by gathering more knowledge and reviewing these results.

The Delphi study is a proven method for eliciting consensus and dissent among experts and identifying potential achievements and developments in the future (Dalkey and Helmer, 1963). A key benefit of the method is that experts around the world can be involved. This was not fully achieved with our sample. We involved experts from different nations, but not from all continents and mainly from the European Union and the United States, so the results may have a Eurocentric bias. One explanation could be that the development of these technologies is still lagging in African and Latin American countries because there are other priorities for DRR. Additionally, we only conducted two rounds, since little consensus was found for the different statements. Our findings indicate the diversity of the topic, as even after two rounds there was still little consensus. However, the experts' answers show some tendencies of opinions and needs. This outcome can be explained by the broadness of the topic but also by the sample size and the participants' characteristics, which are two key limitations within this study. The sample was fairly diverse in terms of the specific research fields of seismology, despite a specific target group being formulated for recruitment. This does not, however, delegitimize the results (Hsu and Sandford, 2007), because the diversity of the group can reveal additional tendencies. It seems that, in order to understand the impacts of these technologies, rather than focusing on a common definition, case studies are helpful to understand the impact of using these technologies for society.

The Delphi study is an appropriate tool to explore policy needs. In the two survey rounds, this was achieved both by showing the differences in the understanding of AI for seismology but also by further developing the toolbox and finding more guiding questions to elicit tendencies as to whether a technology actually enhances DRR and safety culture. These policy needs could be fulfilled by applying a standardized tool for the inclusion of societal matters or targeted funding of research on those matters. Additionally, further research should be conducted with case studies on the other technologies, as well as the different pillars of the toolbox, i.e. the societal dimension and the usability. To this end, it would be beneficial to conduct studies that explore both the acceptance and practical utility of the toolbox, thereby gaining a comprehensive understanding of its usability. Further, to advance the toolbox, it must be actively used and applied by professionals and there must be continuous evaluation of how vulnerability and inclusiveness can be addressed in a technologically fast-evolving world.

6 Conclusion

Emerging technologies such as AI, IoT, and remote sensing are applied in many different fields and can support societies in dealing with disasters. So far, research looking at the practical and societal issues related to emerging technologies for DRR has been limited. This study thus iteratively and inductively developed and tested a toolbox for professionals including developers and researchers, allowing them to critically reflect on and assess the practical and societal impacts of a technology in the context of DRR and safety culture. The toolbox empowers professionals to enhance the accessibility and applicability of their technologies, considering also the needs of vulnerable groups, and encourages a shift in technology assessments from the last to the first mile. Consequently, the societal perspective becomes an integral part of all phases, encompassing the design, development, implementation, and deployment of a technology.

Our case study on AI in seismology has illustrated that the developed toolbox can indeed help and motivate scientists and developers to reflect on the societal issues related to their developments in the context of DRR, but reveals that there is a need for more common understanding and definitions of these technologies, in order for them to be discussed among different professionals.

These technologies have been found to have great potential to enhance DRR and safety culture. We therefore encourage professionals and research groups to use the toolbox for their evaluations of emerging technologies and to further adjust it based on new research findings, since it is a rapidly evolving field and the application always depends on the specific cultural contexts.

Acknowledgements

The authors wish to thank Ariane Wenger and Nikolaj Dahmen for their valuable feedback on the questionnaire. The authors also thank the reviewers for providing valuable comments, which strongly improved the article. This research is part of the sCience and human factOr for Resilient sociEty project (CORE), which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101021746. This research has been approved by the ETH Zurich Ethics Commission (EK 2023-N-15).

7 Data and code availability

All the data can be found as a supplement under: https://doi.org/10.3929/ethz-b-000636485.

8 Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported on in this paper.

References

Adeel, A., Gogate, M., Farooq, S., Ieracitano, C., Dashtipour, K., Larijani, H., and Hussain, A. *A Survey on the Role of Wireless Sensor Networks and IoT in Disaster Management*, page 57–66. Springer Singapore, Aug. 2018. doi: 10.1007/978-981-13-0992-2_5.

Aitsi-Selmi, A., Egawa, S., Sasaki, H., Wannous, C., and Murray, V. The Sendai Framework for Disaster Risk Reduction: Renewing the Global Commitment to People's Resilience, Health, and Well-being. *International Journal of Disaster Risk Science*, 6(2): 164–176, June 2015. doi: 10.1007/s13753-015-0050-9.

Alsalat, G. Y., El-Ramly, M., Aly, A., and Said, K. Detection of Mass Panic using Internet of Things and Machine Learning. *International Journal of Advanced Computer Science and Applications*, 9(5), 2018. doi: 10.14569/ijacsa.2018.090542.

Arabameri, A., Saha, S., Mukherjee, K., Blaschke, T., Chen, W., Ngo, P., and Band, S. In *Modeling Spatial Flood using Novel Ensemble Artificial Intelligence Approaches in Northern Iran*. 2020. page 30.

Arshad, B., Ogie, R., Barthelemy, J., Pradhan, B., Verstaevel, N., and Perez, P. Computer Vision and IoT-Based Sensors in Flood Monitoring and Mapping: A Systematic Review. *Sensors*, 19(22):5012, Nov. 2019. doi: 10.3390/s19225012.

Banna, M. H. A., Taher, K. A., Kaiser, M. S., Mahmud, M., Rahman, M. S., Hosen, A. S. M. S., and Cho, G. H. Application of Artificial Intelligence in Predicting Earthquakes: State-of-the-Art and Future Challenges. *IEEE Access*, 8:192880–192923, 2020. doi: 10.1109/access.2020.3029859.

Bello, O. M. and Aina, Y. A. Satellite Remote Sensing as a Tool in Disaster Management and Sustainable Development: Towards a Synergistic Approach. *Procedia - Social and Behavioral Sciences*, 120:365–373, Mar. 2014. doi: 10.1016/j.sbspro.2014.02.114.

Beroza, G. C., Segou, M., and Mostafa Mousavi, S. Machine learning and earthquake forecasting—next steps. *Nature Communications*, 12(1), Aug. 2021. doi: 10.1038/s41467-021-24952-6.

Boyd, D. and Crawford, K. Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication amp; Society*, 15(5):662–679, June 2012. doi: 10.1080/1369118x.2012.678878.

Bushnaq, O. M., Chaaban, A., and Al-Naffouri, T. Y. The Role of UAV-IoT Networks in Future Wildfire Detection. *IEEE Internet of Things Journal*, 8(23):16984–16999, Dec. 2021. doi: 10.1109/jiot.2021.3077593.

Cao, Y., Boruff, B. J., and McNeill, I. M. The smoke is rising but where is the fire? Exploring effective online map design for wildfire warnings. *Natural Hazards*, 88(3):1473–1501, May 2017. doi: 10.1007/s11069-017-2929-9.

Connell, B., Jones, M., Mace, R., Mueller, J., Mullik, A., Ostroff, E., Sanford, J., Steinfeld, E., Story, M., and Vanderheiden, G. *The principles of universal design—Version 2.0—4/1/97. The center for universal design, N.C.* State University, 1997.

Costache, R. and Tien Bui, D. Spatial prediction of flood potential using new ensembles of bivariate statistics and artificial intelligence: A case study at the Putna river catchment of Romania. *Science of The Total Environment*, 691:1098–1118, Nov. 2019. doi: 10.1016/j.scitotenv.2019.07.197.

Crawford, K. and Finn, M. The limits of crisis data: analytical and ethical challenges of using social and mobile data to understand disasters. *GeoJournal*, 80(4):491–502, Nov. 2014. doi: 10.1007/s10708-014-9597-z.

Crenshaw, K. Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. *Stanford Law Review*, 43(6):1241, July 1991. doi: 10.2307/1229039.

Dalkey, N. The Delphi method: An experimental study of group opinion. In *Studies in the quality of life: Delphi and decision-making*. Lexington Books, Lexington, MA, 1972.

Dalkey, N. and Helmer, O. An experimental application of the delphi method to the use of experts. *Management Science*, 9(3): 458–467, 1963.

Dallo, I., Stauffacher, M., and Marti, M. Actionable and understandable? Evidence-based recommendations for the design of (multi-)hazard warning messages. *International Journal of Disaster Risk Reduction*, 74:102917, May 2022. doi: 10.1016/j.ijdrr.2022.102917.

Datta, A., Wu, D. J., Zhu, W., Cai, M., and Ellsworth, W. L. DeepShake: Shaking Intensity Prediction Using Deep Spatiotemporal RNNs for Earthquake Early Warning. *Seismological Research Letters*, 93(3):1636–1649, Mar. 2022. doi: 10.1785/0220210141.

Dong, L. A comprehensive review of earthquake-induced building damage detection with remote sensing techniques. *ISPRS Journal of Photogrammetry and Remote Sensing*, 15, 2013.

Essam, Y., Kumar, P., Ahmed, A. N., Murti, M. A., and El-Shafie, A. Exploring the reliability of different artificial intelligence techniques in predicting earthquake for Malaysia. *Soil Dynamics and Earthquake Engineering*, 147:106826, Aug. 2021. doi: 10.1016/j.soildyn.2021.106826.

European Commission. Project Cycle Management Guidelines, 2004.

Furquim, G., Filho, G., Jalali, R., Pessin, G., Pazzi, R., and Ueyama, J. How to Improve Fault Tolerance in Disaster Predictions: A Case Study about Flash Floods Using IoT. *Sensors*, 18(3):907, Mar. 2018. doi: 10.3390/s18030907.

Gao, I. Using the Social Network Internet of Things to Mitigate Public Mass Shootings. In *2016 IEEE 2nd International Conference on Collaboration and Internet Computing (CIC)*. IEEE, Nov. 2016. doi: 10.1109/cic.2016.073.

Gevaert, C. M., Carman, M., Rosman, B., Georgiadou, Y., and Soden, R. Fairness and accountability of AI in disaster risk management: Opportunities and challenges. *Patterns*, 2(11):100363, Nov. 2021. doi: 10.1016/j.patter.2021.100363.

Gjøsæter, T., Radianti, J., and Chen, W. Universal Design of ICT for Emergency Management from Stakeholders' Perspective: A Systematic Literature Review. *Information Systems Frontiers*, 23 (5):1213–1225, Nov. 2020. doi: 10.1007/s10796-020-10084-7.

Goyal, H. R., Ghanshala, K. K., and Sharma, S. Flash flood risk management modeling in indian cities using IoT based reinforcement learning. *Materials Today: Proceedings*, 46:10533–10538, 2021. doi: 10.1016/j.matpr.2021.01.072.

Greenhalgh, T. and Peacock, R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *BMJ*, 331(7524):1064–1065, Oct. 2005. doi: 10.1136/bmj.38636.593461.68.

Guerrero, J. Firefighters relying on artificial intelligence to fight California wildfires. In *California Dreaming. ABC 7 News*. 2022. https://abc7news.com/california-wildfires-artificial-intelligence-wifire-lab-burnpro-3d/11596528/.

Harasimiuk, D. E. and Braun, T. *Regulating Artificial Intelligence: Binary Ethics and the Law*. Routledge, Feb. 2021. doi: 10.4324/9781003134725.

Harirchian, E., Kumari, V., Jadhav, K., Rasulzade, S., Lahmer, T., and Raj Das, R. A Synthesized Study Based on Machine Learning Approaches for Rapid Classifying Earthquake Damage Grades to RC Buildings. *Applied Sciences*, 11(16):7540, Aug. 2021. doi: 10.3390/app11167540.

Hsu, C.-C. and Sandford, B. A. The Delphi Technique: Making Sense of Consensus. 2007. doi: 10.7275/PDZ9-TH90.

Huot, F., Hu, R., Goyal, N., Sankar, T., Ihme, M., and Chen, Y.-F. Next Day Wildfire Spread: A Machine Learning Dataset to Predict Wildfire Spreading From Remote-Sensing Data. In *IEEE Transactions on Geoscience and Remote Sensing*, volume 60, page 13. 2022. doi: 10.1109/TGRS.2022.3192974.

Hussein, S. Using remote sensing data for predicting potential areas to flash flood hazards and water resources. *Remote Sensing Applications*, 12, 2019. doi: 10.1016/j.rsase.2019.100254.

Iaccarino, A. G., Gueguen, P., Picozzi, M., and Ghimire, S. Earthquake Early Warning System for Structural Drift Prediction Using Machine Learning and Linear Regressors. *Frontiers in Earth Science*, 9, July 2021. doi: 10.3389/feart.2021.666444.

Ionescu, B., Ghenescu, M., Rastoceanu, F., Roman, R., and Buric, M. Artificial Intelligence Fights Crime and Terrorism at a New Level. *IEEE MultiMedia*, 27(2):55–61, Apr. 2020. doi: 10.1109/mmul.2020.2994403.

IPCC. *Intergovernmental Panel on Climate Change. Summary for Policymakers. Synthesis Report of the IPCC Sixth Assessment Report (AR6)*. 2023. doi: 10.59327/ipcc/ar6-9789291691647.

Ismail-Zadeh, A. T., Cutter, S. L., Takeuchi, K., and Paton, D. Forging a paradigm shift in disaster science. *Natural Hazards*, 86(2): 969–988, Dec. 2016. doi: 10.1007/s11069-016-2726-x.

ISO 22395. Security and resilience—Community resilience—Guidelines for supporting vulnerable persons in an emergency. 2018.

ITU. International Communications Union. GET Background document - Emergency Telecommunications - Disruptive technologies and their use in disaster risk reduction and management. 2019. https://www.itu.int/en/ITU-D/Emergency-Telecommunications/Pages/Events/2019/GET-2019/Disruptive-technologies-and-their-use-in-disaster-risk-reduction-and-management.aspx.

Izumi, T., Shaw, R., Djalante, R., Ishiwatari, M., and Komino, T. Disaster risk reduction and innovations. *Progress in Disaster Science*, 2:100033, July 2019. doi: 10.1016/j.pdisas.2019.100033.

Kaku, K. Satellite remote sensing for disaster management support: A holistic and staged approach based on case studies in Sentinel Asia. *International Journal of Disaster Risk Reduction*, 33:417–432, Feb. 2019. doi: 10.1016/j.ijdrr.2018.09.015.

Kankanamge, N., Yigitcanlar, T., and Goonetilleke, A. Public perceptions on artificial intelligence driven disaster management: Evidence from Sydney, Melbourne and Brisbane. *Telematics and Informatics*, 65:101729, Dec. 2021. doi: 10.1016/j.tele.2021.101729.

Katz, J. and Aspden, P. Motivations for and barriers to Internet usage: Results of a national public opinion. *Internet Research: Electronic Networking Applications and Policy*, 7, Number 3, 1997.

Kaur, H. and Sood, S. K. Fog-assisted IoT-enabled scalable network infrastructure for wildfire surveillance. *Journal of Network and Computer Applications*, 144:171–183, Oct. 2019. doi: 10.1016/j.jnca.2019.07.005.

Kelley, P. G., Yang, Y., Heldreth, C., Moessner, C., Sedley, A., Kramm, A., Newman, D. T., and Woodruff, A. Exciting, Useful, Worrying, Futuristic. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '21. ACM, July 2021. doi: 10.1145/3461702.3462605.

Khan, T., Alam, M., Shahid, Z., and Ahmed, S. Artificial Intelligence based Multi-modal sensing for flash flood investigation. *Applied Sciences*, 6, 2018.

Lucivero, F., Swierstra, T., and Boenink, M. Assessing Expectations: Towards a Toolbox for an Ethics of Emerging Technologies. *NanoEthics*, 5(2):129–141, July 2011. doi: 10.1007/s11569-

011-0119-x.

Majumdar, S., Kumar Pani, S., Kumar Singh, S., Garg, L., Bilas Pachori, R., and Zhang, X. The Role of Remote Sensing and GIS in Military Strategy to Prevent Terror Attacks. In *Intelligent Data Analytics for Terror Threat Prediction: Architectures, Methodologies, Techniques and Applications*, page 79–94. Scrivener Publishing.

Marhain, S., Ahmed, A. N., Murti, M. A., Kumar, P., and El-Shafie, A. Investigating the application of artificial intelligence for earthquake prediction in Terengganu. *Natural Hazards*, 108(1): 977–999, Apr. 2021. doi: 10.1007/s11069-021-04716-7.

Marshall, T. M. Risk perception and safety culture: Tools for improving the implementation of disaster risk reduction strategies. *International Journal of Disaster Risk Reduction*, 47:101557, Aug. 2020. doi: 10.1016/j.ijdrr.2020.101557.

Meier, M., Kodera, Y., Böse, M., Chung, A., Hoshiba, M., Cochran, E., Minson, S., Hauksson, E., and Heaton, T. How Often Can Earthquake Early Warning Systems Alert Sites With High-Intensity Ground Motion? *Journal of Geophysical Research: Solid Earth*, 125(2), Jan. 2020. doi: 10.1029/2019jb017718.

Mishra, A. Remote sensing of extreme flash floods over two southern states of India during North-East monsoon season of 2020. *Natural Hazards*, 6, 2021. doi: 10.1007/s11069-021-04631-x.

Mitra, P., Ray, R., Chatterjee, R., Basu, R., Saha, P., Raha, S., Barman, R., Patra, S., Biswas, S. S., and Saha, S. Flood forecasting using Internet of things and artificial neural networks. In *2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*. IEEE, Oct. 2016. doi: 10.1109/iemcon.2016.7746363.

Mousavi, S. M. and Beroza, G. C. Deep-learning seismology. *Science*, 377(6607), Aug. 2022. doi: 10.1126/science.abm4470.

Mousavi, S. M. and Beroza, G. C. Machine Learning in Earthquake Seismology. *Annual Review of Earth and Planetary Sciences*, 51(1):105–129, May 2023. doi: 10.1146/annurev-earth-071822-100323.

Novellino, A., Jordan, C., Ager, G., Bateson, L., Fleming, C., and Confuorto, P. *Remote Sensing for Natural or Man-Made Disasters and Environmental Changes*, page 23–31. Springer Singapore, Aug. 2018. doi: 10.1007/978-981-13-0992-2_3.

Ogie, R. I., Rho, J. C., and Clarke, R. J. Artificial Intelligence in Disaster Risk Communication: A Systematic Literature Review. In *2018 5th International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)*. IEEE, Dec. 2018. doi: 10.1109/ict-dm.2018.8636380.

Petersen, L., Havarneanu, G., McCrone, N., and Markarian, G. In *Proceedings of the International ISCRAM Conference*. University of Nebraska at Omaha (USA), May 2023. doi: 10.59297/hvjn6870.

Pohl, C. Delphi poll, 2020. doi: 10.5281/ZENODO.3716976.

Rathje, E. M. and Franke, K. Remote sensing for geotechnical earthquake reconnaissance. *Soil Dynamics and Earthquake Engineering*, 91:304–316, Dec. 2016. doi: 10.1016/j.soildyn.2016.09.016.

Ray, P. P., Mukherjee, M., and Shu, L. Internet of Things for Disaster Management: State-of-the-Art and Prospects. *IEEE Access*, 5: 18818–18835, 2017. doi: 10.1109/access.2017.2752174.

Rieland, R. Can Artificial Intelligence Help Stop School Shootings?, 2018. https://www.smithsonianmag.com/innovation/can-artificial-intelligence-help-stop-school-shootings-180969288/.

Sakurai, M. and Shaw, R. *Emerging Technologies for Disaster Resilience: Practical Cases and Theories*. Springer Singapore, 2021. doi: 10.1007/978-981-16-0360-0.

Schubert, K. and Klein, M. *Das Politiklexikon*. Bundeszentrale für politische Bildung, Lizenzausgabe Bonn, 7. aktualisiert und erweiter. aufl edition, 2020.

Seydoux, L., Balestriero, R., Poli, P., Hoop, M. d., Campillo, M., and Baraniuk, R. Clustering earthquake signals and background noises in continuous seismic data with unsupervised deep learning. *Nature Communications*, 11(1), Aug. 2020. doi: 10.1038/s41467-020-17841-x.

Shaw, R. Thirty Years of Science, Technology, and Academia in Disaster Risk Reduction and Emerging Responsibilities. *International Journal of Disaster Risk Science*, 11(4):414–425, Apr. 2020. doi: 10.1007/s13753-020-00264-z.

Singer, N. Schools Are Spending Billions on High-Tech Defense for Mass Shootings, 2022. https://www.nytimes.com/2022/06/26/business/school-safety-technology.html.

Staniforth, A. and Akhgar, B. *Harnessing the Power of Big Data to Counter International Terrorism*, page 23–38. Elsevier, 2015. doi: 10.1016/b978-0-12-801967-2.00003-3.

Steyaert, J. and Gould, N. Social Work and the Changing Face of the Digital Divide. *British Journal of Social Work*, 39(4):740–753, Feb. 2009. doi: 10.1093/bjsw/bcp022.

Stojadinović, Z., Kovačević, M., Marinković, D., and Stojadinović, B. Rapid earthquake loss assessment based on machine learning and representative sampling. *Earthquake Spectra*, 38(1): 152–177, Sept. 2021. doi: 10.1177/87552930211042393.

Sun, W., Bocchini, P., and Davison, B. D. Applications of artificial intelligence for disaster management. *Natural Hazards*, 103(3): 2631–2689, July 2020. doi: 10.1007/s11069-020-04124-3.

Surya, L. Fighting fire with AI: Using deep learning to help predict wildfires in the US, 2020.

Taale, A., Ventura, C. E., and Marti, J. On the feasibility of IoT-based smart meters for earthquake early warning. *Earthquake Spectra*, 37(3):2066–2083, Jan. 2021. doi: 10.1177/8755293020981964.

UNDRR. *United Nations Office for Disaster Risk Reduction (Ed). Our world at risk: Transforming governance for a resilient future*. 2022.

U.N.D.R.R. Disaster Risk Reduction, 2024. https://www.undrr.org/terminology/disaster-risk-reduction.

Verma, S., Kaur, S., Rawat, D. B., Xi, C., Alex, L. T., and Zaman Jhanjhi, N. Intelligent Framework Using IoT-Based WSNs for Wildfire Detection. *IEEE Access*, 9:48185–48196, 2021. doi: 10.1109/access.2021.3060549.

Vickery, J. Using an intersectional approach to advance understanding of homeless persons' vulnerability to disaster. *Environmental Sociology*, 4(1):136–147, Dec. 2017. doi: 10.1080/23251042.2017.1408549.

Vogel, C., Zwolinsky, S., Griffiths, C., Hobbs, M., Henderson, E., and Wilkins, E. A Delphi study to build consensus on the definition and use of big data in obesity research. *International Journal of Obesity*, 43(12):2573–2586, Jan. 2019. doi: 10.1038/s41366-018-0313-9.

Wu, A., Lee, J., Khan, I., and Kwon, Y.-W. CrowdQuake+: Data-driven Earthquake Early Warning via IoT and Deep Learning. In *2021 IEEE International Conference on Big Data (Big Data)*. IEEE, Dec. 2021. doi: 10.1109/bigdata52589.2021.9671971.

Zambrano, A., Perez, I., Palau, C., and Esteve, M. Technologies of Internet of Things applied to an Earthquake Early Warning System. *Future Generation Computer Systems*, 75:206–215, Oct. 2017. doi: 10.1016/j.future.2016.10.009.

Zhao, X., Lovreglio, R., Kuligowski, E., and Nilsson, D. Using Artificial Intelligence for Safe and Effective Wildfire Evacuations. *Fire Technology*, 57(2):483–485, Apr. 2020. doi: 10.1007/s10694-020-00979-x.

doi:10.26443/seismica.v3i1.1142

SEISMICA

Realtime Selection of Optimal Source Parameters Using Ground Motion Envelopes

Dario Jozinović * [1], **John Clinton** [1], **Frédérick Massin** [1], **Maren Böse** [1], **Carlo Cauzzi** [1]

[1]Swiss Seismological Service, ETH Zurich, Zurich, Switzerland

Author contributions: *Conceptualization*: JC, DJ. *Data Curation*: DJ, FM, CC. *Methodology*: DJ, JC, FM, MB. *Formal Analysis*: DJ. *Investigation*: DJ. *Writing - original draft*: DJ, JC, FM, MB. *Writing - Review & Editing*: DJ, JC, FM, MB, CC. *Visualization*: DJ, FrM. *Supervision*: JC.

Abstract It is increasingly common for seismic networks to operate multiple independent automatic algorithms to characterise earthquakes in real-time, such as in earthquake early warning (EEW) or even standard network practice. Commonly used methods to select the best solution at a given time are simple and use ad hoc rules. An absolute measure of how well a solution (event origin and magnitude) matches the observations by the goodness-of-fit between the observed and predicted envelopes is a robust and independent metric to select optimal solutions. We propose such a measure that is calculated as a combination of amplitude and cross-correlation fit. This metric can be used to determine when a preferred solution reaches an appropriate confidence level for alerting, or indeed to compare two (or more) different event characterisations directly. We demonstrate that our approach can also be used to suppress false alarms commonly seen at seismic networks. Tests using the 10 largest earthquakes in Switzerland between 2013 and 2020, and events that caused false alarms demonstrate that our approach can effectively prefer solutions with small errors in location and magnitude, and can clearly identify and discard false origins or incorrect magnitudes, at all time scales, starting with the first event characterisation.

Production Editor:
Gareth Funning
Handling Editor:
Andrea Llenos
Copy & Layout Editor:
Théa Ragon

Received:
October 10, 2023
Accepted:
April 22, 2024
Published:
May 9, 2024

1 Introduction

Over recent decades, earthquake early warning (EEW) algorithms have been continuously developed and EEW systems have become operational in many regions around the world (e.g. Cremen and Galasso, 2020; Allen and Melgar, 2019; Clinton et al., 2016). The goal of EEW is to rapidly estimate developing ground shaking from an ongoing earthquake at a specific location or region, thus providing users with the opportunity to take action before strong ground motions arrive and hence minimise the impact of the shaking.

The Swiss Seismological Service (SED) at Eidgenössische Technische Hochschule (ETH) Zurich has been actively engaged in the development and implementation of EEW algorithms for over a decade (e.g. Massin et al., 2021; Behr et al., 2016), including the Virtual Seismologist (VS; Cua, 2005; Cua and Heaton, 2007) and Finite-Fault Rupture Detector (FinDer; Böse et al., 2012, 2015, 2018, 2023) algorithms. Operationally, these algorithms are integrated as modules within the earthquake monitoring platform SeisComP - a technical framework named the ETHZ-SED SeisComP EEW system (ESE; Massin et al., 2021). Both of these algorithms estimate the earthquake source parameters (location and magnitude) which are then used to estimate ground shaking at a set of locations or a region. In the current implementation, the VS algorithm provides a predictive magnitude, based on the recorded amplitudes and rapid point-source earthquake location provided through existing SeisComP modules. FinDer provides an estimate of the best-fitting line-source model by comparing the peak ground acceleration (PGA) values at stations to a set of pre-generated templates. Both algorithms provide independent estimates of the earthquake source parameters. Differences in their performance during an earthquake can arise from the properties of the seismic source (magnitude, source complexity); the network geometry; the data quality (e.g. both algorithms are sensitive to spikes and metadata errors), and contemporary seismicity. A system building on multiple independent algorithms, like ESE, is more robust and tolerant to failure (Massin et al., 2021).

Like the majority of seismic networks, the SED also uses the SeisComP framework for routine automated monitoring, and also operates a suite of automatic detectors (simple STA/LTA and post-detectors) and locators (scautoloc and scanloc) that target different types of seismicity and use different velocity models. Further, as more and more stations trigger, each 'pipeline' produces updated origins. This means that for an ongoing, even moderate event, the automatic system provides a highly dynamic output with many tens of different origins. Providing solutions from a few stations allows small events to be identified and the earliest solutions for large events. Allowing frequent updates as more phase picks arrive means better accuracy can be achieved as the energy from the seismic event progresses across the seismic network

However, having multiple source parameter solutions requires a method capable of preferring or combining

*Corresponding author: dario.jozinovic@sed.ethz.ch

them. At the SED, we currently use a location score based only on origin parameters (such as number of picks, azimuthal gap, etc.), that does not take magnitude into account. Other approaches are available, for example by simply using a weighted average of the solution parameters (Kohler et al., 2020). Minson et al. (2017) propose a more complex approach in which a constrained least squares fit between observed and predicted acceleration waveform envelopes from all the stations in the network is calculated. This is done for each algorithm solution and used to estimate the relative probability for single solutions or their combinations being correct, with the sum of the probabilities summing to 1. Furthermore, a 'no event' solution, for which the predicted envelopes are equal to zero, is used to obtain the probability of no earthquake existing, corresponding to a false alarm. The method allows the preference or combination of the ground motion predictions according to the relative probabilities assigned. However, it does not provide a measure for the quality of a solution, only its relative quality compared to other possible solutions.

Motivated by the approach of Minson et al. (2017), in this study we also compare the observed and predicted envelopes. We employ the same envelope functions (Cua, 2005) with appropriate adjustments to Switzerland (see the Electronic Supplement to this article). However, instead of calculating the *relative* probabilities of solutions from the given algorithms being correct, we provide an *absolute* goodness-of-fit measure associated with each source parameter estimate, which can then be used to select the preferred solution (i.e. the location, magnitude and origin time) and to check if this solution reaches a pre-defined threshold for alerting. Also, instead of a least squares fit between the observed and predicted envelopes, our approach uses the combination of amplitude fit and cross-correlation between the envelopes, requiring that amplitude and shape fit well but allowing for some timing error, for example in the velocity model. We also do not set the noise level of the predicted envelopes to zero but to the median noise level at the stations used in this study.

Our approach can be implemented in EEW systems as well as in general seismic observatory practice. It can be used to compare multiple available automated seismic solutions, irrespective of the source algorithm and model, and to provide a preferred origin and magnitude. Crucially, it can also be used to compare both point-source (as provided by standard SeisComP locators) and finite-source (as provided by FinDer) solutions, since it can support different distance metrics, including hypocentral or rupture distances.

We test our method on a set of earthquakes and false alarms that occurred in real-time processing at the SED since 2013. The algorithm can successfully prefer source parameter estimates that are close to the network solution with both early solutions including only seconds of data at closest stations, as well as using the full data from a large network. We find it is particularly effective in suppressing processing blunders from significant errors in automated locations or significantly elevated magnitudes that are regular issues in seismic network monitoring.

2 Methods

Our proposed method is based on the comparison of observed and predicted velocity waveform envelopes (Fig. 1) at a set of seismic stations. We obtain the observed earthquake envelopes (following the approach of Cua, 2005), computed by the *sceewenv* SeisComP module (Massin et al., 2021), which provides continuous realtime streams of envelope values (Behr et al., 2016). Envelope computation in *sceewenv* involves the following sequential steps: correct for the gain and baseline offset and check for a clipped signal (neglecting the sensor if has a saturated signal); compute the root-mean-square of the two horizontal components to obtain a single combined horizontal component, integrate to velocity (if needed); apply a 4th order Butterworth high pass filter with a corner frequency of 3 s; and, compute the maximum absolute amplitude within 1 s intervals.

We calculate the predicted envelopes following the Cua (2005) relationship, using magnitude, hypocentral distance and site class (either "rock" or "soil"). To attribute the site class to the stations used in this study, when available, we used the EC8 ground types (Eurocode 8, 2005) available from the SED stations website (Swiss Seismological Service At ETH Zurich, 1983). EC8 ground types are categorised as rock for EC8 ground types A and B) or soil for all other EC8 ground types and as default in the absence of EC8 type. The relationship then outputs P and S waveform envelopes, which start at earthquake origin time with a specified duration that matches the time window of the available station waveform. The Cua (2005) predicted envelopes were calibrated using data from southern California, which consisted of about 30,000 records (vertical and horizontal acceleration, velocity, and displacement) from 70 southern California earthquakes ($2 \leq M \leq 7.3$) recorded within 200 km from the earthquake source. However, for the subset of Swiss earthquakes we used, we observed (Figure S7 in the electronic supplement) that these predicted envelopes often do not fit the observed shaking well and visual checks showed a systematic overpredicting of the observed shaking. Therefore we decided to scale the predicted envelopes using the GMM developed by Cauzzi et al. (2015) for Switzerland (see also Edwards and Fäh, 2013). This approach reduced the difference in peaks between observed and predicted envelopes (Figure S7 in the electronic supplement). While this significantly improved the overall envelope fit, we still found that the maxima of the P-waves (especially at the closest stations) were higher in the observed data. We then further adapted the predicted envelopes using a station-specific S-P scaling and multiplication of the P-wave amplitude with an ad-hoc scaling factor. Further details are given in the Electronic Supplement.

We then compute the goodness-of-fit, G, of the observed and predicted (parts of the) waveform envelopes for each station S at time t as

$$\mathrm{G(S,t)} = 100 \cdot \sqrt{A(S,t) \cdot C(S,t)} \quad (1)$$

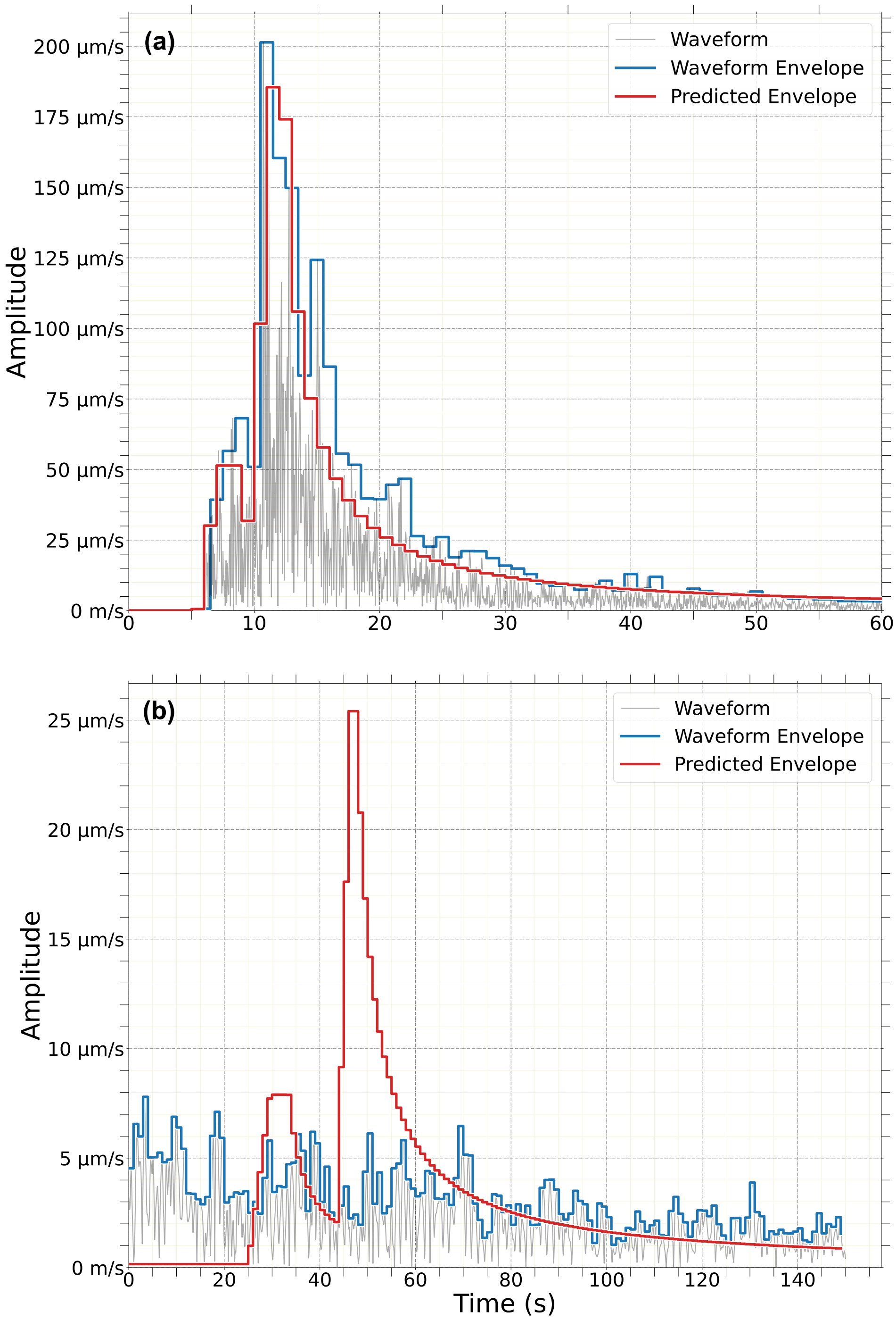


Figure 1 Comparison between true (with accompanying waveform) and predicted envelopes for (a) an earthquake and (b) a false alarm. The subplot a) shows the comparison for an event of M 3.5 at a station 31 km from the hypocentre. The subplot b) shows the comparison for a false M 3.5 at 151 km hypocentral distance, caused by a teleseismic earthquake (M 8.2 in Mexico)

where *A(S, t)* is the amplitude fit, and *C(S, t)* is the normalised zero-shift cross-correlation between the observed and predicted envelopes that start at time t_0 and end at time t, where t_0 can be an arbitrarily defined envelope start time (e.g. earthquake origin time, P-arrival at the closest station, etc.). The amplitude fit, *A(S, t)*, is calculated as

$$\mathrm{A(S,t)} = 1 - \left(\frac{o(S,t) - m(S,t)}{o(S,t) + m(S,t)}\right)^2 \tag{2}$$

where *o(S,t)* and *m(S,t)* are the peak amplitudes of the observed and predicted envelope at time t, respectively. We decided to compare only the maxima of the observed and predicted envelopes as we are modelling the difference in envelope shapes using the cross-correlation fit, *C(S, t)*, calculated as

$$\mathrm{C(S,t)} = \frac{\sum_{i=1}^{n} O(S,t)_i M(S,t)_{(i)}}{\sqrt{\sum_{i=1}^{n} O(S,t)_j^2}\sqrt{\sum_{i=1}^{n} M(S,t)_i^2}} \tag{3}$$

where *O(S, t)* and *M(S,t)* are observed and predicted envelopes at time t, respectively, and n is the number of samples (seconds) in the envelope. While using *C(S, t)* increases the processing time of our algorithm (which

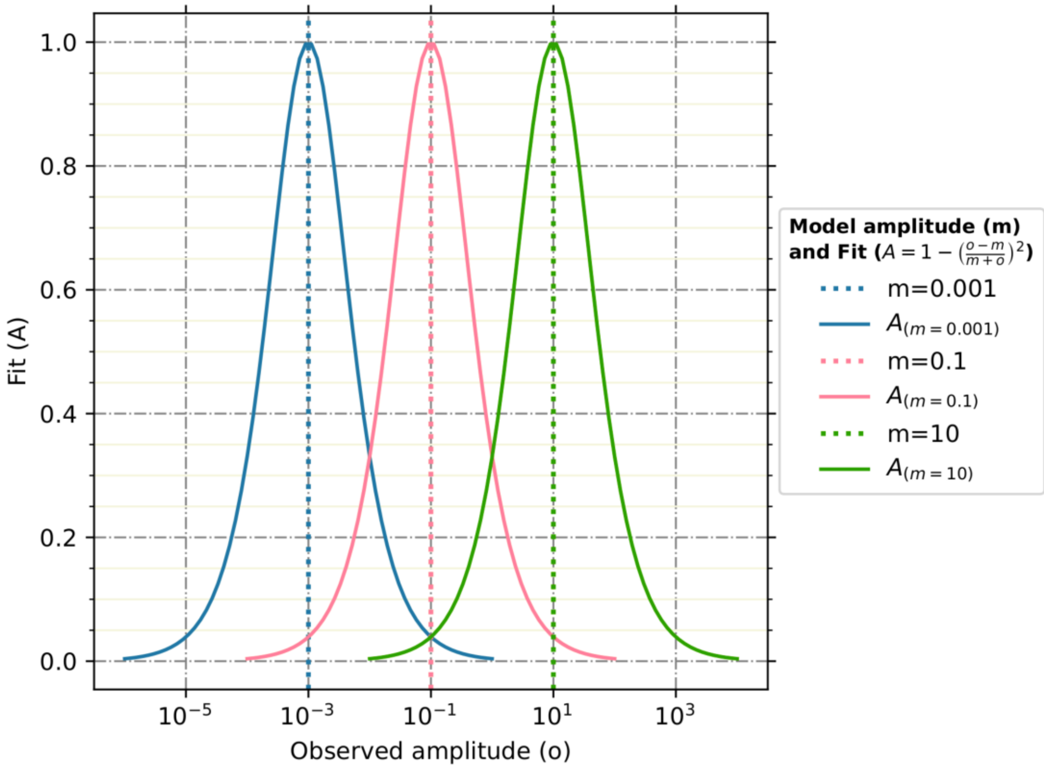


Figure 2 The amplitude fit function A(S, t) in equation 2. The curves show the value of the fit, depending on the difference between the observed (o) and predicted (m) peak values. The predicted values are fixed (dotted vertical line) for 3 different values (3 colours). The amplitude fit A is shown for different observed amplitudes for each of the 3 fixed predicted values. Note that the shape of A is scale-invariant on a log scale.

still remains insignificant), it allows us to address possible (unrealistically) good amplitude fits for certain combinations of both wrong magnitude and distance that can produce similar amplitude maxima as the observed data (e.g. for an M 5.5 earthquake at 50 km distance the predicted PGV is 0.0054 m/s and for an M 5 earthquake at 25 km distance the predicted PGV is 0.0055 m/s), but can be easily discriminated by the envelope shape.

The functional form of $A(S, t)$ in equation 2 was chosen because of its symmetric fit with exponential decay (Fig. 2) as a function of the difference between $o(S,t)$ and $m(S,t)$. Furthermore, we opted to use the functional form of $A(S, t)$ in equation 2 over a least-squares-fit as used in Minson et al. (2017) because the value of $A(S,t)$ is bounded between 0 and 1 and depends on relative differences in predicted and observed amplitudes rather than absolute values (i.e. larger amplitudes do not affect the fit disproportionately). Crucially, this produces a bounded absolute fit measure that is independent of earthquake size, allowing us to prefer the best magnitude and location estimate (although it systematically penalises weaker motions - for example for an observed ground motion of 2 *mm/s*, predicted ground motions 1 *mm/s* and 4 *mm/s* both provide the same $A(S,t)$ value).

Finally, the mean goodness-of-fit across all included stations at time t is calculated and used as the measure of the goodness-of-fit for the given source parameters. The same procedure is applied again for both existing and new source parameter estimates as more and more data arrives from already included and newly introduced stations - resulting in a goodness-of-fit metric that evolves over time. We then choose the solution with the highest mean goodness-of-fit, i.e. that best fits the observed ground motions, as our preferred solution.

In practice, at any given time t, we only use stations where the predicted or observed earthquake ground-motion envelopes are non-zero - that is those stations within a certain distance, dependent on time t, from the hypocenter (or rupture plane). The duration of the predicted envelope matches the available data from the observed stations at time t. In this manner, the method can account for differing data latencies from seismic stations if operated in real-time. A further benefit of this approach is that it reduces processing time and does not allow the fit between observed and predicted noise (which is hard to model, and is station or sensor specific) to affect the final goodness-of-fit measure.

3 Data

We evaluate the performance of the proposed algorithm in two separate tests using data collected by the Swiss seismic network. This first is a retrospective analysis using recent moderate earthquakes. The second test analyses a set of significant false earthquake alerts that were produced in recent years. We explore both (1) how sensitive our method is at identifying differences in location and magnitudes, and (2) how effective it is at identifying false alarms. All tests are performed using different window lengths after the P-wave arrival at the clos-

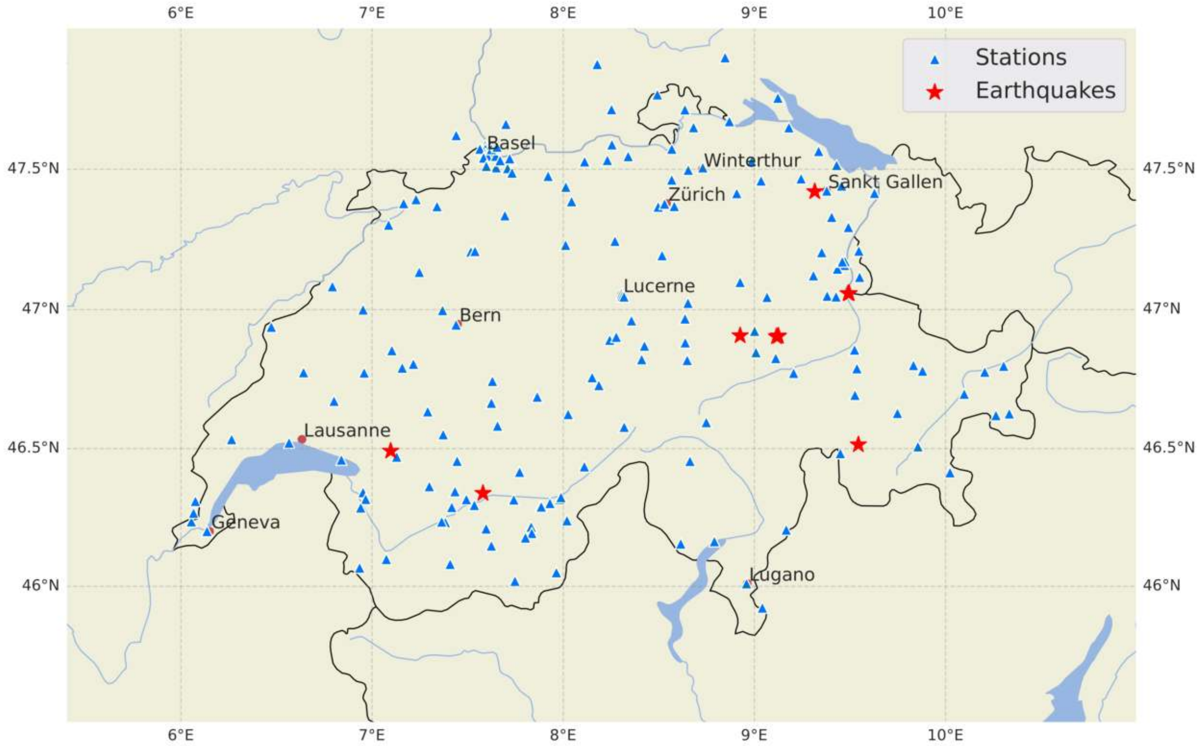


Figure 3 Map of the stations (network codes CH, C4) and the 10 Swiss earthquakes (Table 1 for details of the events) used in the analysis. Some earthquakes have very similar locations, so their markers overlap on the map.

Date (UTC)	Latitude (°N)	Longitude (°E)	Magnitude (Mlh)	Depth (km)
2020-11-10 T 12:53	46.9	9.12	3.9	1.7
2020-10-25 T 19:43	46.9	9.12	4.3	1.4
2020-10-25 T 19:35	46.91	9.12	3.6	1.4
2017-07-01 T 08:10	46.49	7.1	4.3	4.3
2017-03-06 T 20:12	46.91	8.93	4.6	4.2
2016-10-24 T 14:44	46.34	7.58	4.1	8.2
2016-10-07 T 07:27	46.51	9.54	3.8	10
2013-12-27 T 07:08	47.06	9.5	3.7	6.2
2013-12-12 T 00:59	47.06	9.49	4.1	5.9
2013-07-20 T 03:30	47.42	9.32	3.5	4.5

Table 1 Earthquakes used in this study.

est station, ranging from t=1 s to t=40 s. We conduct our tests using the earthquakes and stations located in Switzerland, a sufficiently small region for which 40 s long waveforms provide enough input data to our algorithm. The sample of recent earthquakes includes the ten largest events that occurred in Switzerland between 2013 and 2020 (Table 1; Fig. 3), with events magnitudes ranging from 3.5 to 4.6 and depths of 1.4 to 10 km.

The sample of 20 false alarms (Table 2) comprises real events (quarry blasts and regional or teleseismic earthquakes) and non-existing events (i.e. noise bursts), that were assigned wrong (or any in case of non-existing events) locations and magnitudes due to combinations of false triggers during routine monitoring. These events typically were released to the public (e.g. on social media platforms) by the alerting system at the Swiss Seismic Network. Alerts are released for automatic solutions with M>2.5 and an epicenter lying inside or close to Switzerland that reach the quality threshold based on the SED ’location score’. Currently, solutions (including false alarms) are generated using the scautopick module in SeisComP, using a minimum of 6 associated picks. The SED ’location score’ (Diehl et al., 2015) takes into account the distribution of pick residuals, location azimuthal gap, location RMS, and the number of arrivals used for location.

We obtain the earthquake and station metadata from the ETHZ dataset in the SeisBench package (Woollam et al., 2022). The ETHZ dataset is derived from 1) the National Earthquake Catalogue of Switzerland (see Data and Resources section) for earthquake information; and 2) the ETH EIDA node for seismic waveform data and metadata. Seismic stations used are from network code CH (Swiss Seismological Service At ETH Zurich, 1983) and C4 (C.E.R.N., 2016), comprising 178 stations. See Figure 3 for stations and events. The false

Real event	Date	False Latitude (°N)	False Longitude (°E)	False magnitude	False depth (km)
Quarry Blast (M=1.6) 46 km from false alarm location	12.06.2019	47.09	7.53	2.8	39.9
Quarry Blast (M=1.5) 140 km from false alarm location	26.08.2020	47.14	9.01	2.8	10
Quarry Blast (M=2) 0 km from false location	1.10.2020	47.53	8.18	2.5	6.4
Quarry Blast (M=1.7) 38 km from false location	5.11.2020	47.31	7.81	2.6	37.7
Quarry Blast (M=0.9) 105 km from false location	01.07.2022	47.17	7.19	2.9	10
M6.4 in Crete, Greece, 1584 km from false location	12.10.2013	46.72	11.46	3.6	12.3
M5.2 in Greece, 1274 km from false location	30.03.2019	46.37	11.21	3.1	27.8
M5.2 in Greece, 2042 km from false location	30.01.2020	46.83	8.94	2.6	76.2
M5.3 in Crete, Greece, 1993 km from false location	23.05.2020.	46.53	8.35	2.6	54.4
M5.4 event in Greece, 1664 km from false location	18.12.2021	46.06	8.27	2.6	27.4
M5.1 event in southern Italy, 916 km from false location	31.10.2022	46	8.19	2.8	3.6
M7.9 in Papua New Guinea, 119° from false location	22.01.2017	46.97	10.31	3.3	61.2
M8.2 in Chiapas, Mexico, 90° from false location	08.09.2017	46.38	11.78	3.5	22.8
M6.8 in Fiji, 151° from false location	18.11.2018	46.78	8.84	2.8	58.1
M7.1 in Anchorage, Alaska, 71° from false location	30.11.2018	46.35	10.18	3.6	3
M6.7 in Fiji, 152° from false location	01.09.2019.	46.69	9.57	2.7	36.5
Non-existing event	11.01.2019	44.82	8.13	3.7	27.6
Non-existing event	05.07.2019	46.54	9.38	3.2	10
M3.6 event, 115 km from false location	25.10.2020	47.48	7.85	2.7	5.9
M3.7 event in Albstadt, Germany, 248 km from false location	21.03.2021	46.51	6.9	2.5	28.5

Table 2 False alarms used in this study.

alarms are representative examples collected from the operational experience during past years.

The observed earthquake envelopes have been calculated by the *sceewenv* module in SeisComP, as described in the Methods 2 section. We calculate the predicted envelopes following the Cua (2005) relationship (details explained in the Electronic Supplement). We set the noise in the envelope templates equal to 10^{-7} m/s (corresponding to the median noise level at the stations used). When calculating the goodness-of-fit for an earthquake we choose the P-arrival at the closest station as the start time t_0 of the envelopes.

4 Results and Discussion

In order to evaluate the performance of our proposed algorithm in the scope of EEW, we calculate the mean goodness-of-fit for 10 large Swiss earthquakes (described in the data section) using a wide range of perturbations of the true source parameters, at different times after the event can be first identified. This test assesses the sensitivity of the algorithm to errors in the source parameters. Specifically, we perturb the catalogue magnitude (varying between -1 and 1 in increments of 0.1 magnitude units, with additional gross perturbations of -1.5 and 1.3 magnitude units; the perturbations are corresponding to $M_{predicted} - M_{true}$) and catalogue epicentral location (by distances of 0, 1, 3, 5, 10, 15, 20, 25, 30, 100, and 150 km), and calculate the goodness-of-fit for all possible combinations of these perturbations. We do not perturb the origin time, though origin time errors can have a strong impact on the goodness of fit. Algorithms that do not provide strong constraints on the origin times can be penalised by this metric, and further work is required to address this. We conduct the same analysis for various input window lengths, from windows that end 2s after the data has reached the 1st station, to 40s after, spanning the entire time window from the earliest possible EEW solution using the minimum available data from the fewest stations, to a final automatic location based on the full waveform from the entire network. To perturb the epicentral location, we randomly vary the latitude and longitude of the epicentre in a manner that satisfies the required distance perturbation.

Figure 4 and Figure 5 show the results of all the permuted magnitude-distance pairs for input windows that end 4 and 20 s after the first P-arrival, respectively. Figures S1 (2 s), S2 (3 s), S3 (7 s), and S4 (40 s) in the Electronic Supplement show the same calculations for a wider range of time windows. Our method would provide robust discrimination to larger errors in magnitude and location already at 4 s after the first P-arrival time (Figure 4). Even at this early stage, the method prefers a solution with correct source parameters (i.e. the one with zero perturbation in magnitude and distance). While smaller errors in distance (less than 10 km) and magnitude (less than 0.5 magnitude units) around the true solution also have relatively high fit values (close to or higher than 55), we can see that the fit falls significantly as the perturbation in both magnitude and distance increases. Increasing the window length

4 s window

Distance error (km) (columns); **Magnitude error** (rows)

	0	1	3	5	10	15	20	25	30	100	150
-1.50	25.46	24.73	23.15	22.03	20.59	14.93	11.76	9.17	5.68	0.00	0.00
-1.0	38.81	37.79	34.96	33.55	28.38	20.92	14.35	11.80	7.07	0.44	0.00
-0.9	41.95	40.84	38.87	36.71	29.85	23.40	16.54	10.77	7.65	0.40	0.23
-0.8	45.14	44.03	40.61	39.02	31.93	24.32	18.12	12.48	7.87	0.00	0.00
-0.7	48.29	47.08	43.40	42.63	33.83	25.10	19.32	12.87	7.59	0.08	0.00
-0.6	51.61	49.81	45.94	43.50	38.08	27.81	20.04	14.72	9.60	0.06	0.00
-0.5	*55.08	53.01	47.69	44.67	39.38	28.78	18.66	14.99	8.58	0.01	0.00
-0.4	*57.90	*55.05	50.29	48.99	39.83	25.87	21.68	12.83	10.45	0.40	0.00
-0.3	*59.82	*56.79	51.58	49.09	37.76	30.36	22.02	13.03	8.73	0.08	0.18
-0.2	*61.28	*58.10	53.00	49.19	39.96	29.70	19.53	16.89	8.66	0.40	0.04
-0.1	*62.29	*58.92	53.72	50.47	41.97	30.38	21.64	16.53	8.14	0.09	0.03
0.0	*62.81	*59.10	53.88	48.83	41.37	26.79	19.79	13.68	10.95	0.33	0.02
0.1	*61.52	*58.46	*57.21	49.30	39.60	25.72	21.99	17.23	7.99	0.26	0.00
0.2	*59.48	*57.33	52.48	48.58	37.78	28.22	19.27	14.97	9.22	0.49	0.05
0.3	*56.67	*55.43	*56.39	47.08	33.60	30.09	21.91	16.96	7.73	0.17	0.02
0.4	53.39	52.82	51.57	47.64	34.95	25.72	23.46	18.96	8.38	0.15	0.05
0.5	49.83	50.00	47.53	45.24	32.80	27.01	22.79	13.59	9.74	0.05	0.05
0.6	46.41	47.24	45.42	40.49	34.16	23.43	21.78	14.58	10.19	0.15	0.00
0.7	42.98	44.56	45.63	39.36	31.07	24.67	19.90	15.12	12.08	0.05	0.20
0.8	39.85	41.48	40.81	38.15	30.62	25.57	19.92	15.67	9.67	0.01	0.01
0.9	36.86	38.51	38.05	35.15	30.23	21.79	20.32	17.34	9.61	0.03	0.04
1.0	33.98	35.69	37.97	29.79	28.62	20.05	18.78	12.96	12.22	0.22	0.00
1.30	27.04	28.81	29.33	26.29	24.62	18.34	15.51	12.28	8.72	0.03	0.00

Figure 4 Average goodness-of-fit for different magnitude-epicentral distance perturbation pairs averaged taken over all 10 events in Table 1 using an input time window ending 4 seconds after the P-arrival at the closest station. The columns and rows show the errors in the source location (km) and magnitude, respectively. The small star in front of a number is used to mark the goodness-of-fit value higher than 55.

20 s window

Distance error (km) (columns); **Magnitude error** (rows)

	0	1	3	5	10	15	20	25	30	100	150
-1.50	19.56	19.44	19.35	18.62	15.66	13.11	12.10	9.74	8.82	4.41	1.13
-1.0	32.63	32.24	31.58	30.81	27.16	22.70	19.30	16.89	15.84	7.15	1.82
-0.9	36.19	35.77	35.35	33.60	30.67	25.40	21.67	19.21	19.49	8.64	3.80
-0.8	40.11	39.63	39.35	38.38	33.92	28.16	24.44	21.30	19.37	7.74	2.64
-0.7	44.33	43.88	43.07	42.65	37.15	31.61	26.56	24.07	22.55	11.34	5.46
-0.6	48.88	48.27	47.73	46.48	40.69	34.81	31.87	26.85	24.79	11.35	4.04
-0.5	53.70	53.18	52.72	51.87	45.75	39.11	33.15	29.61	26.92	10.15	4.50
-0.4	*58.58	*58.16	*57.44	*55.98	49.17	42.06	36.26	32.15	29.42	13.38	5.76
-0.3	*63.37	*62.89	*62.20	*60.47	53.55	45.07	38.24	35.39	31.82	13.99	4.41
-0.2	*67.15	*66.65	*65.68	*63.88	*57.12	47.76	39.83	35.17	32.01	14.52	8.45
-0.1	*69.23	*68.75	*67.75	*65.26	*58.41	49.57	41.85	38.89	33.79	14.43	5.09
0.0	*70.03	*69.56	*68.33	*65.77	*58.27	48.64	43.03	39.79	35.42	18.27	9.46
0.1	*69.12	*68.71	*67.22	*65.74	*58.89	50.84	42.46	40.72	37.32	19.53	6.44
0.2	*66.46	*66.14	*64.68	*62.93	*57.11	48.80	44.61	37.27	35.07	19.47	8.56
0.3	*62.98	*62.88	*61.50	*60.68	*55.36	47.19	43.83	40.27	34.04	18.81	9.40
0.4	*58.41	*58.23	*57.58	*55.42	51.87	45.51	42.38	39.63	37.13	21.80	8.41
0.5	53.76	53.70	52.41	52.12	48.79	40.29	39.65	33.03	36.97	21.50	11.13
0.6	49.26	49.22	48.23	47.04	41.82	40.78	34.03	36.72	30.09	19.75	7.35
0.7	44.82	44.92	44.15	42.91	38.39	37.43	32.13	34.86	31.62	20.76	15.09
0.8	40.68	40.96	39.62	40.54	36.34	32.68	31.94	30.08	33.02	18.38	10.42
0.9	37.05	37.19	36.57	36.20	34.41	31.48	30.99	25.97	29.05	19.16	8.40
1.0	33.67	33.89	34.17	32.21	32.13	26.84	26.90	26.64	27.73	18.71	12.86
1.30	26.09	26.33	26.64	24.89	24.46	22.29	22.60	19.76	22.00	15.75	9.71

Figure 5 Same as Figure 4 for an input time window ending 20 seconds after the first P-arrival at the closest station

(Figure 5) improves the absolute fit for solutions with small magnitude and distance perturbations. Furthermore, there is an improvement in sensitivity to magnitude perturbations (i.e. the fit for larger magnitude errors is decreasing compared to the shorter time window). However, in terms of location error, the sensitivity deteriorates slightly. This can be explained by the inclusion of more distant stations into the goodness-of-fit calculation. Figure S5 in the electronic supplement presents the same results as Figure 5, after 20 s, except only stations up to 50km from the epicenter are used. For more distant stations, a small change in hypocentral distance does not have a large effect on envelope amplitude. Furthermore, for solutions with significant distance errors, the observed increase in the goodness-of-fit value at longer time windows is also a consequence of the inclusion into the GOF calculation of 1) more distant stations for which the predicted amplitudes are often close to the noise level (Figure S6a), and 2) the stations for which the true and false epicentre can be at a

Earthquakes

Time after P-arrival at the closest station	2020-11-10 M=3.9 T(1) = 1.8 s T(4)-T(1) = 1.4 s	2020-10-25 M=4.3 T(1) = 1.9 s T(4)-T(1) = 1.3 s	2020-10-25 M=3.6 T(1) = 1.9 s T(4)-T(1) = 1.3 s	2017-07-01 M=4.3 T(1) = 1.2 s T(4)-T(1) = 2.6 s	2017-03-06 M=4.6 T(1) = 1.6 s T(4)-T(1) = 2.4 s
1s	0.00	NaN	NaN	60.33	0.00
2s	77.89	84.12	60.34	92.44	52.61
3s	56.90	51.34	61.37	25.39	54.44
4s	59.30	65.76	60.29	75.15	53.56
5s	68.85	63.08	73.88	53.44	44.63
7s	69.19	73.95	65.68	67.69	53.68
9s	68.94	74.36	64.62	69.38	59.33
12s	64.89	66.28	59.87	66.46	56.84
15s	74.56	70.48	72.62	69.61	62.83
19s	74.54	69.65	71.86	73.99	71.85
25s	76.06	72.56	72.56	75.69	74.43
30s	75.14	73.95	73.09	74.49	74.34
40s	72.77	71.75	71.34	74.81	75.29

Time after P-arrival at the closest station	2016-10-24 M=4.1 T(1) = 2.1 s T(4)-T(1) = 1 s	2016-10-07 M=3.8 T(1) = 3.1 s T(4)-T(1) = 2 s	2013-12-27 M=3.7 T(1) = 2.3 s T(4)-T(1) = 0.3 s	2013-12-12 M=4.1 T(1) = 2.4 s T(4)-T(1) = 2.3 s	2013-07-20 M=3.5 T(1) = 1.4 s T(4)-T(1) = 4 s
1s	9.98	75.44	38.71	24.71	57.70
2s	32.95	61.64	75.20	70.38	53.40
3s	30.35	80.09	70.79	72.29	59.23
4s	50.68	75.68	68.38	55.37	63.94
5s	45.65	48.19	66.51	54.54	48.56
7s	48.01	51.39	63.40	58.03	68.76
9s	52.82	49.15	67.44	59.87	65.91
12s	50.50	53.35	60.25	60.04	69.27
15s	55.52	55.24	70.19	67.18	70.49
19s	58.13	58.06	70.93	67.06	76.00
25s	61.11	64.66	72.80	67.67	75.66
30s	64.21	65.87	72.35	68.95	75.72
40s	64.97	64.60	71.87	70.92	74.90

Figure 6 The variation of the goodness-of-fit for the individual earthquakes in Table 1, over time, assuming the correct location and magnitude. The rows show the window length after the P-arrival at the closest station, and the column headers show the date, the magnitude of the earthquake, the travel times for the first station T(1), and the difference between the travel times at the first and the fourth stations T(1)-T(4). The hatched cells show the times when the P-waves would not yet have reached 4 stations - i.e. before an EEW alert could be released. The small star in the right bottom corner of the cell is used to mark the goodness-of-fit value higher than 55.

similar distance (Figure S6b), e.g. a station halfway between the true and false epicentre.

For the purpose of the current study, we consider 55 as the goodness-of-fit threshold which we would use to define a solution acceptable. However, a more in-depth analysis in real-time testing (as discussed later in the text) is required to define a threshold which would minimize false alerts while allowing for selection of good solutions.

Figures S1, S2, S3 and S4, presenting different time windows, show similar patterns as in Figures 4 and 5. However, at very short time windows, the goodness of fit does not always increase with time - at 2 s fit values at and close to the correct solution are higher than at 3 s, 4 s or 7s. This is a consequence of having a smaller number of stations in the goodness-of-fit calculation (for the shorter time windows) which have a strong individual influence on the final goodness-of-fit metric.

Figure 6 summarises the evolution in goodness of fit over time for the correct or perfect solution (i.e. using the network derived location and magnitude, so distance and magnitude error are zero) for each of the 10 largest individual events. The same general variations seen for the combined events persist for the individual events. At 1 s, solutions are often missing (none of the stations have their amplitudes above the amplitude threshold) or have a very poor fit. Note in the column header, the time after the origin time of the 1st and 4th pick is provided for each event. This time difference is a proxy for the very earliest time an EEW solution could be available using a network-type EEW approach like for VS or FinDer. For the dense Swiss network, this value ranges from 0.3 to 4 s - and in many cases, a first solution would be available only 2 s after the 1st arrival. In Figure 6, shaded cells indicate when the goodness of fit would be computed before a 4-station EEW solution is possible. In the first few seconds, we also observe strong goodness-of-fit variations for individual earthquakes, which is mostly a consequence of the wrong travel times predicted through the envelope prediction relationship (we did not adapt the travel times to Switzerland). The goodness-of-fit improves for all of the individual earthquakes over time, and for the majority of them converges to a similar value when longer time windows are used. If we take 55 as a goodness-of-fit threshold value we can see that the algorithm would prefer the correct solution from the time the first EEW alert has been issued for 7 events, while for 3 events it would take 9 and 12 s to reach the threshold.

Figures 7 to 10 show the result of the goodness-of-fit for the false alarm origins listed in Table 2. The method again shows good performance, it would allow for false alarm discrimination in practically all cases. The method provides extremely low goodness-of-fit values (mostly zero) for all the examples of quarry blasts causing false alarms except one (Figure 7). The example with the high goodness-of-fit value - just reaching the 55 threshold for some window lengths - actually had the correct location, though the magnitude was overestimated by 0.5 magnitude units, making it by far the 'least wrong' of all the examples. The goodness of fit values over time for these examples are consistent with the previous Figures 4, 5 and S1-S4. Extremely low goodness-of-fit results are also observed for false alarms caused by events close to Swiss borders and for false alarms without a specific event being the cause of the false alarm (both types in Figure 10). The algorithm can also clearly discriminate against the false alarms caused by teleseismic and regional earthquakes (Figures 8 and 9). We can also see that the goodness-of-fit tends to increase with increasing time window length for some of the false alarms. As noted before, this results from the inclusion of more distant stations into the calculation for which the predicted envelope amplitudes are close to the noise level. It should be noted that a goodness-of-fit value of zero was assigned in cases when no stations reached the required threshold level (explained at the beginning of the chapter). The predicted envelopes do not include event type. However, the efficacy of the method on different event types is demonstrated in Figures 7 to 10.

The method is shown to be highly effective, with favourable results observed almost immediately (2 s) after the P-wave arrival at the first station. This suggests our goodness-of-fit metric can be used to select the preferred of multiple EEW solutions in real-time. The method not only ranks the EEW solutions but also provides a measure of their absolute quality, which can be used to decide whether any of the solutions is acceptable for emitting an alert. The goodness-of-fit does increase for all solutions (even very wrong ones) with the increase of the input window length (i.e. adding more stations). This, however, could be tackled by adjusting the goodness-of-fit threshold with the increase of the input window length, or by weighting the stations according to the distance from the (predicted) epicentre.

The algorithm successfully penalises significant errors in both magnitude and location. Since the approach is based on matching the actually observed ground motions, it allows for the integration of location and magnitude accuracy estimates into a single quality value. Hence, it provides an independent and fair comparison of very different algorithms, including those that produce a point-source solution with those that produce line-source solutions (for which we could use e.g. the Joyner-Boore distance metric). The results also show that the method effectively suppresses false alarms. We expect that this approach, if integrated into real-time monitoring frameworks, will surpass the performance of the traditional metrics that combine simple parameters (e.g. the number of picks, RMS, azimuthal gap, etc.) and hence can replace them.

A challenge with this method is that the background station noise can be above the predicted ground motions, especially when analysing signals from small earthquakes. High background noise can come from anthropogenic sources or indeed sensor noise if the sensor quality is limited, for example from MEMS accelerometers. In these cases, the final goodness-of-fit value can be dominated by the noise rather than the signal. In this study, we simplify the predicted noise modelling by using the median noise level of the stations used in the analysis, which is a reasonable assumption given the overall high quality of the Swiss network. However, for more heterogeneous networks it may be needed to make the noise modelling more station- or sensor-specific.

In actual network operations, we would restrict the station selection to only stations with predicted P-arrivals. However, in our study using perturbed distance and magnitude errors, we could not select the stations based on the predicted P-arrival times as we did not have the difference in origin times for the different distance perturbations. Thus we only select stations close to the real (catalogue) and predicted hypocenter using the previously described amplitude threshold. Applying this selection allows the cross-correlation fit to decrease the influence of noisy stations on the final goodness-of-fit.

Future improvement in the station selection procedure could come from weighting the stations so as to ensure the most relevant stations have the highest influence on the final goodness-of-fit. This could be achieved by 1) weighting the stations according to their epicentral distance which would reduce the effect of distant stations that have the predicted ground motions close to the noise level and, in a reasonably homogeneously spaced network, are usually more numerous than the more important stations near the epicenter; and 2) down-weighting stations located in spatial clusters as to limit the influence of areas with a high density of stations. Weighting, however, could also have a negative effect (e.g. for large location errors, weighting by distance from the predicted epicenter could downweight important stations near the real epicenter) and requires a detailed analysis of the whole network and the individual stations when implementing the algorithm. We expect that adopting a weighting procedure will allow us to reduce the effect of distant stations increasing of overall goodness-of-fit for wrong solutions as can be observed between Figures 4 and 5. This will also allow

Time after P-arrival at the closest station	01-07-2022, M(F)=2.9, M(T)=0.9, De=105 km	05-11-2020, M(F)=2.6, M(T)=1.7, De=38 km	01-10-2020, M(F)=2.5, M(T)=2, De=0 km	26-08-2020, M(F)=2.8, M(T)=1.5, De=140 km	12-06-2019, M(F)=2.8, M(T)=1.6, De=46 km
1 s	0.00	0.00	41.15	0.00	0.00
2 s	0.00	0.00	27.97	0.00	0.00
3 s	0.00	0.00	40.03	0.00	0.00
5 s	0.00	0.00	54.65	0.00	0.00
7 s	0.00	0.00	54.81	0.00	0.00
10 s	0.48	0.00	* 55.40	0.00	0.00
13 s	0.00	0.00	* 55.89	0.00	7.36
17 s	2.52	7.88	42.27	0.00	14.58
23 s	3.14	7.58	42.60	3.44	10.05
28 s	2.70	5.55	42.79	4.03	4.97
38 s	2.79	3.98	46.15	11.16	5.04

Figure 7 The goodness-of-fit for the false alarms in Table 2. The column header shows the date of the event (as a proxy for ID); the false magnitude *M(F)*; the true magnitude *M(T)*; and distance error *De* if caused by a real event. The first columns shows the length of the input time window (in seconds) after the first theoretical P-arrival (using the false location). The small star in front of a number is used to mark the goodness-of-fit value higher than 55. The results are grouped according to the source of the false alarm; here, for **false alarms caused by quarry blasts**.

Time after P-arrival at the closest station	31-10-2022, M M(F)=2.8, M(T)=5.1, De=916 km	18-12-2021, M(F)=2.6, M(T)=5.4, De=1664 km	23-05-2020, M(F)=2.6, M(T)=5.3, De=1993 km	30-01-2020, M(F)=2.6, M(T)=5.2, De=2042 km	30-03-2019, M(F)=3.1, M(T)=5.3, De=1274 km	12-10-2013, M(F)=3.6, M(T)=6.4, De=1584 km
1 s	0.00	0.00	0.00	0.00	0.00	0.00
2 s	0.00	0.00	0.00	0.00	0.00	0.00
3 s	0.00	0.00	0.00	0.00	0.00	0.00
5 s	0.00	0.00	0.00	0.00	0.00	0.00
7 s	0.00	0.00	0.00	0.00	0.00	0.00
10 s	0.00	0.00	0.00	0.00	0.00	0.00
13 s	17.00	0.00	0.00	0.00	0.00	0.00
17 s	15.88	0.00	0.00	0.00	0.00	15.40
23 s	12.16	0.00	0.00	0.00	0.00	14.33
28 s	10.03	0.00	0.00	0.00	0.00	9.48
38 s	7.93	21.73	0.00	0.00	0.00	0.00

Figure 8 Same as Fig. 7, for false alarms caused by regional earthquakes in the Mediterranean.

Time after P-arrival at the closest station	01-09-2019, M(F)=2.7, M(T)=6.7, De=152°	30-11-2018, M(F)=3.6, M(T)=7.1, De=71°	18-11-2018, M(F)=2.8, M(T)=6.8, De=151°	08-09-2017, M(F)=3.5, M(T)=8.2, De=90°	22-1-2017, M(F)=3.3, M(T)=7.9, De=119°
1 s	0.00	0.00	0.00	0.00	0.00
2 s	0.00	0.00	0.00	0.00	0.00
3 s	0.00	16.60	0.00	0.00	0.00
5 s	0.00	9.81	0.00	0.00	0.00
7 s	0.00	5.80	0.00	0.00	0.00
10 s	0.00	2.98	0.00	0.00	0.00
13 s	0.00	3.45	0.00	3.24	0.00
17 s	0.00	2.79	0.00	5.07	0.00
23 s	0.00	11.62	0.00	7.34	0.00
28 s	0.00	19.23	0.00	6.87	0.00
38 s	0.00	31.87	0.00	5.18	0.00

Figure 9 Same as Fig. 7, for false alarms caused by teleseisms.

Time after P-arrival at the closest station	21-03-2021, M(F)=2.5, M(T)=3.7, De=248 km	25-10-2020, M(F)=2.7, M(T)=3.6, De=115 km	05-07-2019, M(F)=3.2, False triggers from a non-existing event	11-01-2019, M(F)=3.7, False triggers from a non-existing event
1 s	0.00	0.00	0.00	0.00
2 s	0.00	10.39	0.00	0.00
3 s	0.00	10.05	15.41	0.00
5 s	0.00	12.28	0.00	0.00
7 s	0.00	12.42	0.00	0.00
10 s	0.00	16.43	3.87	0.00
13 s	0.00	16.84	2.61	0.00
17 s	0.00	20.02	5.19	4.52
23 s	4.05	22.74	16.40	7.22
28 s	4.09	24.60	15.33	5.72
38 s	4.07	27.25	20.21	10.70

Figure 10 Same as Fig. 7, for other false alarms. The first 2 are mislocations close to Switzerland from larger regional events, the second 2 are from non-existing events.

us to more precisely select the goodness-of-fit threshold which we would use to accept a solution.

The processing time of the algorithm (on a personal laptop - Lenovo ThinkPad T14 Gen 2a) was on average 0.65 s per earthquake (tested on the 10 Swiss earthquakes) without significant variation when using different window lengths. This means that the processing time is dominated by loading the observed and predicted envelopes from disk - the calculation of the goodness-of-fit took on average 0.003 s when the envelopes were loaded into the memory. The main improvement in the processing time can then be achieved by loading the envelope data faster (e.g. loading only the envelopes from the triggered stations - in the experiment we loaded the envelopes from all the stations).

Our approach is applicable to any monitoring system, though it relies on having an appropriate set of predicted envelopes for the seismicity being monitored. For Switzerland, as described in the Supplement, we used the original envelope prediction relationship developed by Cua (2005) that was developed using data from Southern California, with modifications to adapt it for Switzerland. Direct application of the method to other regions would likely require customising the envelopes for the specific region or accepting a reduced performance in terms of goodness-of-fit values. Furthermore, it is unclear how well the envelope prediction relationships apply to large (bigger than M 6.5) earthquakes, which could affect the goodness-of-fit values for those events. Some preliminary tests on this topic have been done (Yamada and Heaton, 2008), but more extensive testing is required to confirm these results. To make the method more general, our next steps in improving the method will include developing a more general envelope prediction method developed on a global earthquake dataset with a significant representation of large events (ideally uniform across magnitudes).

The tests in this study were not done in real-time, i.e. we did not account for actual station latencies. On the other hand, we were using only the stations from the CH and C4 networks, meaning that more data could be available from other networks. We were also missing the real-time trigger information and had to rely on an amplitude threshold as a proxy for triggered stations which could allow non-relevant stations to enter the final goodness-of-fit value. Having real-time information about the event origin time (for the correct or false solutions) will actually improve the performance, as it allows us to select stations based on expected P-arrivals, i.e. only those that are relevant. As noted before, a station weighting procedure will be explored to increase the effect of relevant stations on the overall goodness-of-fit. We relied on the travel times calculated using the original envelope prediction relationship, which resulted in wrong start (onset) times for some of the envelopes. Finally, we expect that the errors in timing/signal quality/metadata could strongly affect the results of the method. The test of the influence of these errors on the results of the method will be made during the real-time implementation of the method, where unplanned errors can occur. Given all the unknowns just described, the real-time implementation of the method will also allow us to understand the performance of the method during times of normal (i.e. low magnitude) seismicity. Therefore, real-time testing of our method is necessary to further confirm it as a practical tool for seismic networks and EEW systems which is the crucial next step in the implementation of the algorithm at SED.

5 Conclusions

We have developed an algorithm that allows the preferred location and magnitude selection for EEW and real-time seismic processing and can be used to suppress false alarms. The algorithm computes a goodness-of-fit between emerging observed velocity waveform envelopes at multiple stations in a seismic network and those predicted by the given origin and magnitude. Our algorithm has been developed and tested on 10 Swiss earthquakes with magnitudes from 3.5 to 4.6, and on 20 events that caused false alarms inside the Swiss monitoring network. Results in this study suggest the proposed algorithm can operate effectively in EEW systems as well as in routine seismic processing. Strong performance is observed for a range of input window lengths, starting from a few seconds after the P-wave arrival at the first station to longer input window lengths, making the algorithm highly suitable for real-time use. The incorporation of the method into a real-time environment brings more challenges beyond just the calculation of the goodness-of-fit. However, the method can bring significant benefits to operational (EEW and earthquake monitoring) systems, justifying the effort needed to implement it. Future improvements will include: improving amplitude fits by re-calibrating the envelope functions using recently collected data, potentially including regionalisation; improving the predicted onset times; and weighting (clusters of) stations (especially at distance).

Data and code availability

The observed envelope data and the envelope templates of (Cua, 2005, , not-adapted to Switzerland) are available at https://zenodo.org/records/10037549, together with the Python code. The station amplification factors, needed for GMM calculations when adapting the predicted envelopes to Switzerland are available at https://stations.seismo.ethz.ch/en/home/ (ETH Zurich, Swiss Seismological Service, 2015). The earthquake and station metadata are available through the ETHZ dataset in SeisBench (Woollam et al., 2022).

Acknowledgements

The Editor wishes to thank two anonymous reviewers for their help in evaluating this article. This study was funded through the Real-time Earthquake Risk Reduction for a Resilient Europe 'RISE' project funded from the European Union's Horizon 2020 research and innovation program under grant agreement No. 821115. Opinions expressed in this paper solely reflect the authors' view; the EU is not responsible for any use that may be made of the information it contains. Contributions from FM are possible thanks to the ATTAC (Alerta Temprana para Terremotos en América Central) project funded by the Swiss Agency for Development and Cooperation (SDC).

Competing interests

There are no competing interests.

References

Allen, R. and Melgar, D. Earthquake early warning: Advances, scientific challenges, and societal needs. *Annual Review of Earth and Planetary Sciences*, 47:361–388, 2019. doi: 10.1146/annurev-earth-053018-060457.

Behr, Y., Clinton, J., Cauzzi, C., Hauksson, E., Jónsdóttir, K., Marius, C., Pinar, A., Salichon, J., and Sokos, E. The Virtual Seismologist in SeisComP3: A new implementation strategy for earthquake early warning algorithms. *Seismological Research Letters*, 87 (2A):363–373, 2016. doi: 10.1785/0220150235.

Böse, M., Heaton, T., and Hauksson, E. Real-time Finite Fault Rupture Detector (FinDer) for large earthquakes. *Geophysical Journal International*, 191(2):803–812, 2012. doi: 10.1111/j.1365-246X.2012.05657.x.

Böse, M., Felizardo, C., and Heaton, T. Finite-Fault Rupture Detector (FinDer): Going real-time in Californian ShakeAlert warning system. *Seismological Research Letters*, 86(6):1692–1704, 2015. doi: 10.1785/0220150154.

Böse, M., Smith, D., Felizardo, C., Meier, M., Heaton, T., and Clinton, J. FinDer v. 2: Improved real-time ground-motion predictions for M2–M9 with seismic finite-source characterization. *Geophysical Journal International*, 212(1):725–742, 2018. doi: 10.1785/0120220183.

Böse, M., Andrews, J., Hartog, R., and Felizardo, C. Performance and Next-Generation Development of the Finite-Fault Rupture Detector (FinDer) within the United States West Coast ShakeAlert Warning System. *Bulletin of the Seismological Society of America*, 2023. doi: 10.1785/0120220183.

Cauzzi, C., Edwards, B., Fäh, D., Clinton, J., Wiemer, S., Kästli, P., Cua, G., and Giardini, D. New predictive equations and site amplification estimates for the next-generation Swiss ShakeMaps. *Geophysical Journal International*, 200(1):421–438, 2015. doi: 10.1093/gji/ggu404.

C.E.R.N. *CERN Seismic Network*. ETH Zurich. Other/Seismic Network, 2016. doi: 10.12686/sed/networks/c4.

Clinton, J., Zollo, A., Marmureanu, A., Zulfikar, C., and Parolai, S. State-of-the art and future of earthquake early warning in the European region. *Bulletin of Earthquake Engineering*, 14: 2441–2458, 2016. doi: 10.1007/s10518-016-9922-7.

Cremen, G. and Galasso, C. Earthquake early warning: Recent advances and perspectives. *Earth-Science Reviews*, 205:103184, 2020. doi: 10.1016/j.earscirev.2020.103184.

Cua, G. *Creating the Virtual Seismologist: developments in ground motion characterization and seismic early warning*. California Institute of Technology, 2005.

Cua, G. and Heaton, T. The Virtual Seismologist (VS) method: A Bayesian approach to earthquake early warning. *Earthquake early warning systems*, page 97–132, 2007. doi: 10.1007/978-3-540-72241-0_7.

Diehl, T., Kästli, P., and Heimers, S. A location quality score - Documentation of the SeisComP3 Implementation for Seattle/Jakarta, Jun 2015. https://gitlab.seismo.ethz.ch/sed-sc3/evscore/.

Edwards, B. and Fäh, D. A Stochastic Ground-Motion Model for Switzerland. *Bulletin of the Seismological Society of America*, 103 (1):78–98,, 2013. doi: 10.1785/0120110331.

ETH Zurich, Swiss Seismological Service. The Site Characterization Database for Seismic Stations in Switzerland, 2015. doi: 10.12686/sed-stationcharacterizationdb.

Eurocode 8. Design of structures for earthquake resistance-part 1: general rules, seismic actions and rules for buildings. *Brussels: European Committee for Standardization*, 2005.

Kohler, M., Smith, D., Andrews, J., Chung, A., Hartog, R., Henson, I., Given, D., Groot, R., and Guiwits, S. Earthquake early warning ShakeAlert 2.0: Public rollout. *Seismological Research Letters*, 91(3):1763–1775, 2020. doi: 10.1785/0220190245.

Massin, F., Clinton, J., and Böse, M. Status of earthquake early warning in Switzerland. *Frontiers in Earth Science*, 9:707654, 2021. doi: 10.3389/feart.2021.707654.

Minson, S., Wu, S., Beck, J., and Heaton, T. Combining multiple earthquake models in real time for earthquake early warning. *Bulletin of the Seismological Society of America*, 107(4): 1868–1882, 2017. doi: 10.1785/0120160331.

Swiss Seismological Service At ETH Zurich. National Seismic Networks of Switzerland. *ETH Zürich. Other/Seismic Network*, 1983. doi: 10.12686/SED/NETWORKS/CH.

Woollam, J., Münchmeyer, J., Tilmann, F., Rietbrock, A., Lange, D., Bornstein, T., Diehl, T., Giunchi, C., Haslinger, F., Jozinović, D., and Michelini, A. SeisBench—A toolbox for machine learning in seismology. *Seismological Society of America*, 93(3):1695–1709, 2022. doi: 10.1785/0220210324.

Yamada, M. and Heaton, T. Real-time estimation of fault rupture extent using envelopes of acceleration. *Bulletin of the Seismological Society of America*, 98(2):607–619, 2008. doi: https://doi.org/10.1785/0120060218.

doi:10.26443/seismica.v3i1.1135

Discontinuous transtensional rupture during the Mw 7.2 1995 Gulf of Aqaba earthquake

Hannes Vasyura-Bathke * [1,2,3], **Andreas Steinberg** [4], **Frank Krüger** [2], **Guangcai Feng** [5], **P. Martin Mai** [1], **Sigurjón Jónsson** [1]

[1]King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia, [2]University of Potsdam, D-14476 Potsdam, Germany, [3]Now at: Helmholtz Centre Potsdam, German Research Centre for Geosciences GFZ, [4]BGR, Federal Institute for Geosciences and Natural Resources, D-30655 Hannover, Germany, [5]Central South University, Changsha, China

Author contributions: *Conceptualization*: H. Vasyura-Bathke, S. Jónsson. *Methodology*: H. Vasyura-Bathke, A. Steinberg. *Software*: H. Vasyura-Bathke, A. Steinberg. *Formal Analysis*: H. Vasyura-Bathke, A. Steinberg, G. Feng. *Writing - original draft*: H. Vasyura-Bathke, A. Steinberg. *Writing - Review & Editing*: H. Vasyura-Bathke, A. Steinberg, F. Krüger, P. M. Mai, S. Jónsson, G. Feng. *Visualization*: H. Vasyura-Bathke. *Supervision*: F. Krüger, P. M. Mai, S. Jónsson. *Funding acquisition*: S. Jónsson.

Abstract The Gulf of Aqaba earthquake occurred on 22 November 1995 in the Northern Red Sea and is the largest instrumentally recorded earthquake in the region to date. The event was extensively studied during the initial years following its occurrence. However, it remained unclear which of the many faults in the gulf were activated during the earthquake. We present results from multi-array back projection that we use to inform Bayesian kinematic rupture models constrained by geodetic and teleseismic data. Our results indicate that most of the moment release was on the Aragonese fault via left-lateral strike slip and shallow normal faulting that may have been dynamically triggered by an early rupture phase on the Arnona fault. We also identified a predominantly normal-fault segment on the eastern shore of the gulf that was activated in the event. We dismiss the previously proposed hypothesis of a co-seismic sub-event on the western shore of the gulf and confirm that observed deformation can be rather attributed to post-seismic activity. In conclusion, the gulf shows many signs of active tectonic extension. Therefore, more events close to the shorelines are to be expected in the future and should be considered when conducting infrastructure projects in the region.

Non-technical summary The 1995 Gulf of Aqaba earthquake was a significant event that has been reexamined using a modern, multifaceted approach. By combining space geodetic satellite data and seismic waveform data, we have gained a more complete understanding of the earthquake, while taking into account potential errors in our analysis. Our results show that during the 1995 earthquake, three faults were activated across distinct fault segments: the Arnona fault in the south, the adjacent Aragonese fault in the north, and an undisclosed fault on the eastern shore of the gulf. These faults exhibited predominantly horizontal motion but also revealed a significant vertical component, underscoring the extension of the gulf. This discovery holds profound implications, particularly given the considerable infrastructure projects currently underway in the Gulf of Aqaba, i.e. within NEOM. In light of these developments, it is evident that earthquake modeling in the region is of paramount importance. The findings from this study underscore the necessity for updated hazard assessments and the establishment of plausible scenarios for potential future earthquakes.

Production Editor: Gareth Funning
Handling Editor: Mathilde Radiguet
Copy & Layout Editor: Théa Ragon

Received: October 20, 2023
Accepted: February 2, 2024
Published: February 19, 2024

1 Introduction

The Gulf of Aqaba is located at the southern end of the Dead Sea fault, which is a left-lateral transform fault system with an estimated average slip rate of $\sim 5 \pm 1$ mm/yr during the Holocene (Le Béon et al., 2012; Lefevre et al., 2018). Geodetic observations show that the current left-lateral interseismic motion in the gulf is similar to that of the Dead Sea fault, with a small amount of opening across the gulf (ArRajehi et al., 2010; Li et al., 2021; Castro-Perdomo et al., 2022; Viltres et al., 2022). This transtensional motion has resulted in a complex tectonic setting of several transform faults and pull-apart basins within the 180-km-long Gulf of Aqaba (Ben-Avraham, 1985; Ribot et al., 2021, Fig.1). The area has also been the seismically most active part of the Dead Sea transform fault with persistent micro-earthquake activity, several seismic swarms, and major events in the past several decades (e.g., Klinger et al., 1999).

The 22 November 1995 (gCMT time 04:15:26.2) Gulf of Aqaba earthquake (Mw 7.2) is the largest instrumentally recorded event in the northern Red Sea and along the entire 1000-km-long Dead Sea transform fault system (Fig. 1). Multiple studies on the earthquake have been published with both point and finite-fault models estimated, either from seismic data (Pinar and Türkelli, 1997; Klinger et al., 1999; Hofstetter et al., 2003) or geodetic data (Klinger et al., 2000; Baer et al., 2001; Shamir et al., 2003; Baer et al., 2008, Supplement Tab. S2, Fig.1). Due to the use of different datasets and the complex tectonic setting, the derived models are diverse and have high epistemic uncertainty, to the point

*Corresponding author: vasbath@gfz-potsdam.de

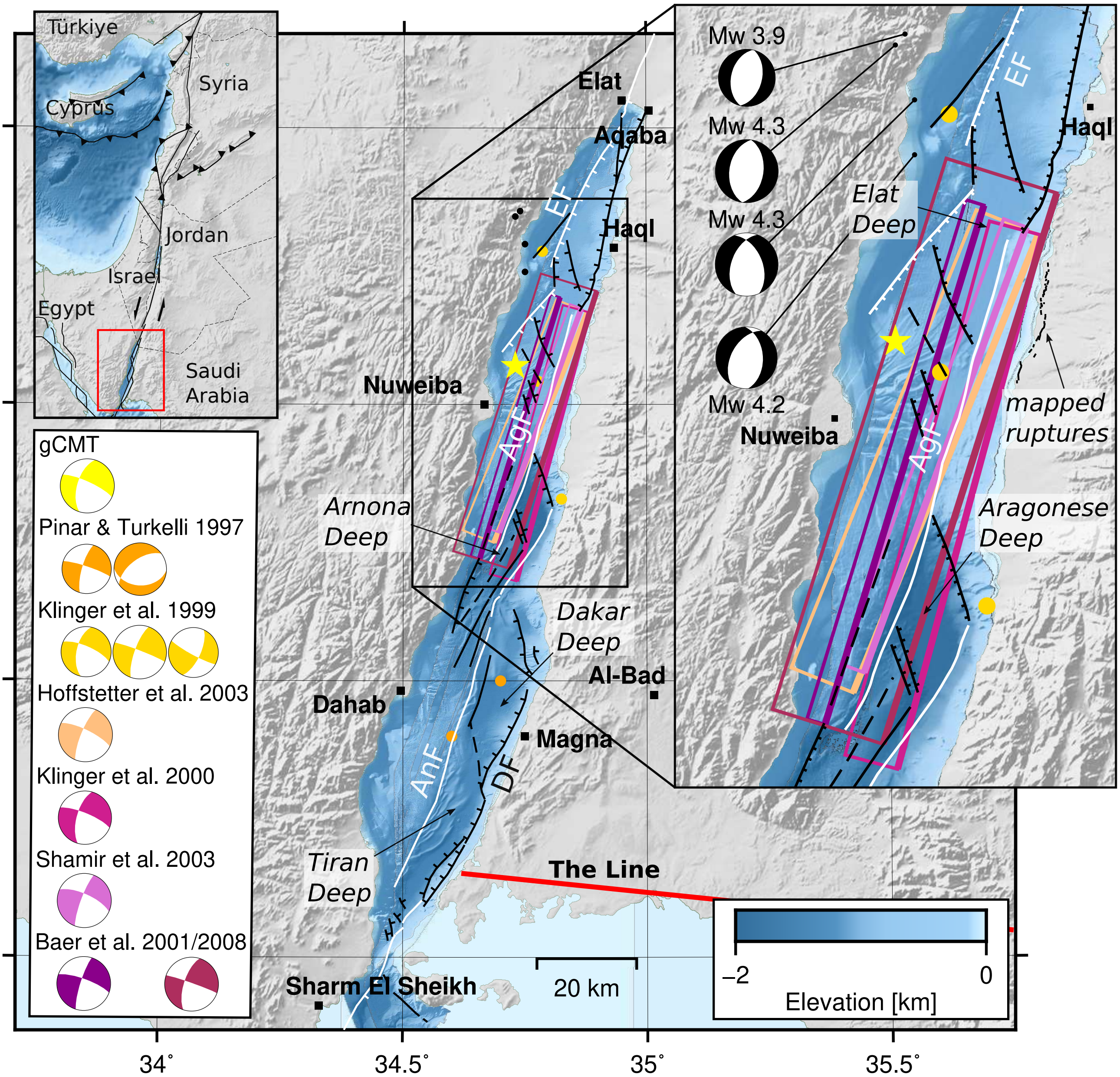


Figure 1 Map of the Gulf of Aqaba and previously estimated finite and point source solutions (from the list compiled in Supplement Tab. S2). The Gulf of Aqaba consists of three pull-apart basins from North to South; the Elat Deep, the Aragonese Deep and the Dakar-Tiran Deep bounded by three main strike-slip faults (white) and normal faults (black) (Ribot et al., 2021). Focal mechanisms of the 1995 event (sizes scaled by magnitude, Supplement Tab. S2) are color coded by the respective location of estimated point (dots) or finite sources (rectangles) in the maps. The upper right inset shows a zoom-in to the epicentral region, where on-land normal faulting (black) attributed to the 1995 event was mapped on the East coast of the Gulf (Lefevre, 2018). The four black focal mechanisms are selected aftershocks to the Gulf of Aqaba earthquake potentially contributing to post-seismic deformation (Hofstetter et al., 2003; Baer et al., 2008). Shaded topography is from SRTM3 (Farr et al., 2007) and bathymetry data is from (Ribot et al., 2021). AgF- Aragonese Fault, AnF- Arnona Fault, EF- Elat Fault, DF- Dakar Fault. The yellow star shows the epicentre of the Gulf of Aqaba earthquake as determined by gCMT. The upper left inset shows the region of interest (red) at the southern Dead Sea transform fault system (modified from Castro-Perdomo et al. (2022)).

that there is no clear consensus on the rupture process of the earthquake. Nonetheless, most of the studies found that the majority of the seismic moment was released on the Aragonese Fault (AgF) in the central part of the gulf (Fig. 1). However, visible complexities in teleseismic broadband waveforms and in surface-displacement maps derived from interferometric synthetic aperture radar (InSAR) suggest a more complex multi-fault rupture in the gulf. Based on the teleseismic waveform data complexity, different number of sub-events and candidates of potentially activated faults have been proposed, which involve the Elat Fault (EF), the Arnona Fault (AnF) and the Dakar Fault (DF), as well as the Aragonese Fault (Pinar and Türkelli, 1997; Klinger et al., 1999; Shamir et al., 2003).

The sparsity of available near-field geodetic data and far-field seismic data makes it difficult to study the details of the earthquake rupture process. The main fault rupture was off-shore within the gulf and InSAR data are, therefore, not available in the near-field of the earthquake (Klinger et al., 2000; Baer et al., 2001, 2008). Also, the roughly N-S striking orientations of the involved faults limit the capability of obtaining the full surface displacement field with InSAR, as radar line-of-sight (LOS) observations are not very sensitive to the predominant North-South coseismic surface displacements. Furthermore, SAR-image acquisitions were infrequent and irregular in this region in the 1990s (Supplement Tab. S1). Finally, notable post-seismic activity, particularly aftershocks, were reported on the Egyptian side of the gulf near the town of Nuweiba (Klinger et al., 1999; Baer et al., 2008). Therefore, the co-seismic displacement field cannot be clearly isolated from secondary deformation processes after the earthquake. These challenges contribute to the uncertainties of estimated geodetic fault-slip models and probably in part explain the large differences between them (Supplement Tab. S2).

Seismic data analysis of the earthquake also faces challenges. At the time of the earthquake, the regional networks of seismic stations were sparse and seismic data were not easily shared between the four countries bordering the gulf (Saudi Arabia, Egypt, Israel, Jordan). Therefore, spatially uneven station geometry was used by the different agencies to locate the aftershock sequence of the earthquake (Hofstetter et al., 2003). Consequently, locations and faulting mechanisms of aftershocks are associated with large uncertainties (Abdel Fattah et al., 1997; Hofstetter et al., 2003).

With independent information on the location of the main fault rupture, identified and mapped based on data from a recent multibeam bathymetric survey (Ribot et al., 2021), with previously unused geodetic data, and applying multi-array teleseismic backprojection, we here derive a refined kinematic finite-fault rupture model for the Gulf of Aqaba earthquake using Bayesian inference combining geodetic and seismic data. This allows us to obtain a clearer picture of the rupture propagation during the earthquake.

2 Data and Methods

We use two main methods to study the rupture evolution of the 1995 Gulf of Aqaba earthquake: multi-array teleseismic backprojection (BP) and Bayesian kinematic earthquake source inference. We use the results obtained from the BP as a-priori information to parameterize the fault geometry and to constrain the parameter solution space for the source inference.

2.1 Multi-array backprojection

We applied multi-array teleseismic BP to image the spatio-temporal evolution of the rupture of the 1995 Aqaba earthquake. In traditional BP seismic recordings of a single array of seismic stations are windowed, aligned and stacked with respect to theoretical traveltimes calculated from a layered Earth structure model for a horizontal grid of potential source locations (Krüger and Ohrnberger, 2005; Ishii et al., 2005). This procedure involves no assumption on the fault geometry and allows space-time imaging of coherent seismic energy radiation (Kiser and Ishii, 2017) through so-called semblance maps. Note that semblance amplitudes have no direct physical relation to the amount of fault slip, but semblance maps indicate where radiated seismic energy of a certain slowness is coherently emitted. Coherent high-frequency seismic energy radiation is expected near the hypocenter in case of an energetic rupture onset and near asperities (patches of high slip), representing (abrupt) changes in earthquake rupture speed (Madariaga, 1977; Spudich and Frazer, 1984; Ide, 2002; Okuwaki and Yagi, 2018). Therefore, mapping the occurrence of high-frequency seismic radiation allows to highlight changes in the fault geometry and to assess the evolution and complexity of the earthquake rupture process independently from the source inversion. Instead of conducting a classical single large-array backprojection, we used multiple small-scale arrays for backprojection. Here, we outline the method only briefly, further details can be found in Steinberg et al. (2022).

We used a multi-array backprojection approach by clustering the available seismic stations using the k-means algorithm (Steinhaus, 1956) into many small virtual arrays (Rössler et al., 2010). We considered all globally available broad-band stations with waveform records between $23°$ and $93°$ distance from the earthquake epicenter, to have all near-field terms attenuated and to avoid triplications. Each virtual array has a unique set of stations and a minimum of 4 stations with a maximum array aperture of 3.5°. For the 1995 Gulf of Aqaba earthquake we ended up with 14 virtual arrays, with the smallest and largest arrays consisting of 4 and 26 stations, respectively (Fig. S1a). We only used the vertical component of velocity waveform records and backprojected with respect to the expected arrival times of the P-phase. Unfortunately, the sparse station configuration did not allow to also use the S-phase. We then calculated the semblance for each virtual array applying a phase-weighted stacking method (Schimmel and Paulssen, 1997). Semblance is a dimensionless mea-

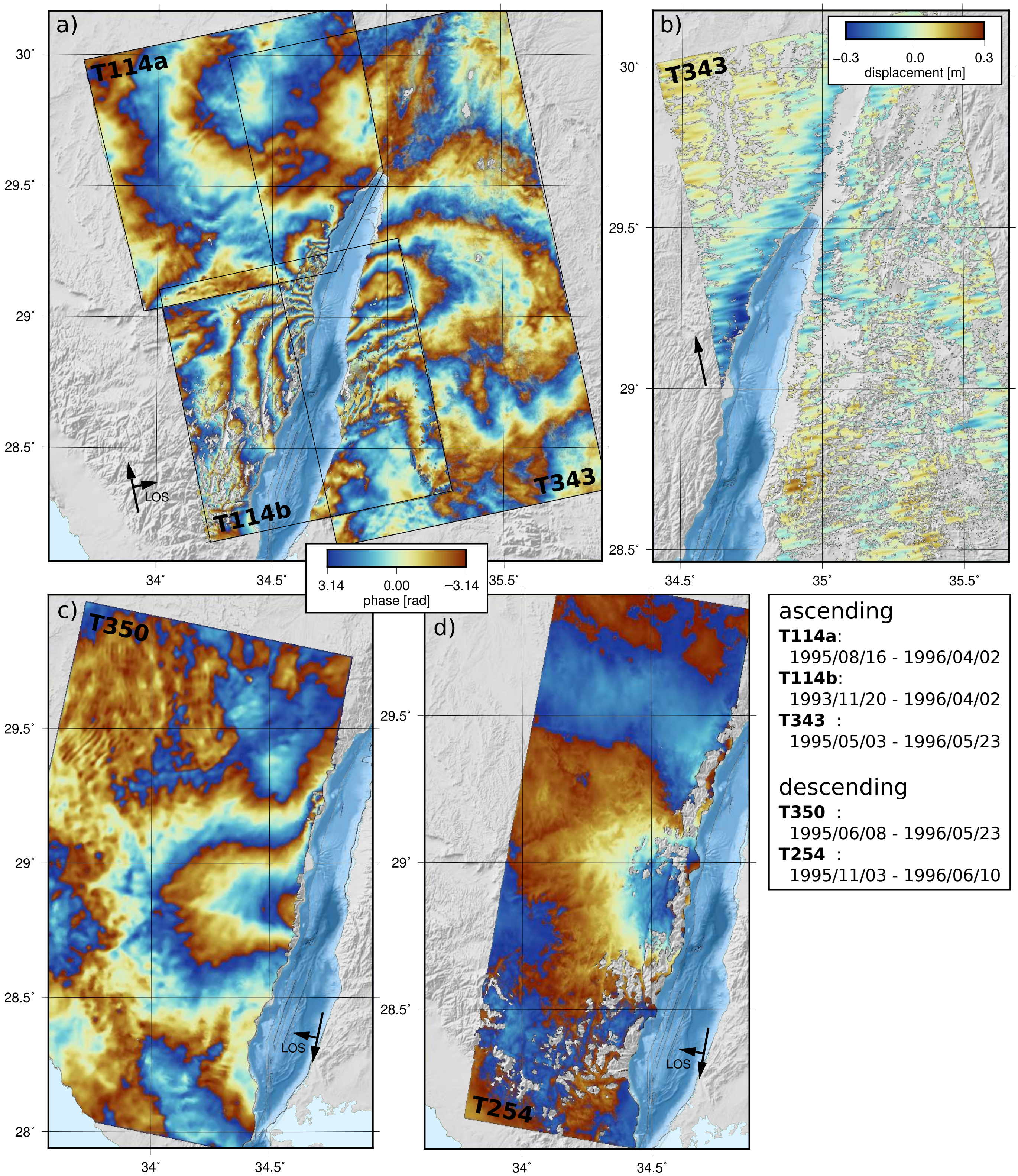


Figure 2 a, c, d) Coseismic interferograms and b) azimuth pixel-offsets derived from SAR data (Supplement Tab. S1) and used for the fault geometry and finite-fault inference. Long and short arrows indicate the flight- and line-of-sight (LOS) directions of the satellite and the radar, respectively. Note that the data of track 254 are L-band data from the JERS-1, whereas the other are C-band data from the ERS-1/2 satellites. Bold black text labels indicate the satellite track numbers of the different datasets. The bottom right panel shows the primary and secondary SAR image acquisition dates for each track (format YYYY/MM/DD).

sure of coherence of waveforms at an array and can be multiplied, similar to a likelihood (Rössler et al., 2010). Therefore, we obtained the multi-array semblance by multiplication of all semblances (Steinberg et al., 2022).

Three-dimensional wave propagation effects that cannot be described by the assumed AK-135 1-D layered Earth structure model (Kennett and Engdahl, 1991) cause shifts in the wave arrival time and these in turn bias the semblance map. To mitigate this effect and to reduce the absolute location error of the multi-array semblance map we estimate empirical travel time shifts and calibrate the waveform data before stacking at each station (Palo et al., 2014; Ishii et al., 2007; Meng et al., 2016; Fan and Shearer, 2017). We estimated these travel time shifts by maximizing the semblance for each virtual array based on a spatially close aftershock to the 1995 Gulf of Aqaba earthquake, the Mw 5.7 earthquake from 23.11.1995, 18:07:17 at $29.333° N, 34.749° E$ (USGS location).

2.1.1 Configuration for the Gulf of Aqaba earthquake

For the Gulf of Aqaba earthquake we backprojected moving time windows of 6s length every 2s on point locations of a horizontal grid with a spacing of 0.05 degree (∼5 km) at a depth of 18 km. We downsampled the seismic recordings to a common 10 Hz and we bandpass filtered the data, above the corner frequency of 0.15 Hz and up to 1.5 Hz. Following this process, we thus focus only on the high-frequency coherent emissions. The waveform records were then stacked for each time-step with respect to the calibrated theoretical arrival time for each considered grid point.

2.1.2 Backprojection uncertainty

To quantify the spatial and temporal accuracy of the multi-array BP results we follow a bootstrapping approach (Wang et al., 2016; Meng et al., 2012) on the semblance calculation at each timestep by perturbing seismic-wave travel times uniform randomly between $\pm 2s$ for each virtual array and additionally between $\pm 0.2s$ for each station. This process allows assessing the influence of wavepath effects on the multi-array semblance. We thus obtain an ensemble of bootstrapped multi-array semblance maps on which quantiles of semblance can be calculated. To document the uncertainties associated with the BP results we display multi-array semblance maps for each timestep, in which semblance maps show all possible coherent seismic-radiation locations for each timestep on the considered grid.

2.2 Bayesian kinematic finite fault inference

We use the Bayesian Earthquake Analysis Tool (BEAT, Vasyura-Bathke et al., 2019) and apply the step-wise inference strategy of Vasyura-Bathke et al. (2020) on surface-displacement maps derived from synthetic aperture radar (SAR) and on broadband teleseismic waveforms to infer distributions of earthquake source parameters that explain the observations within the range of associated uncertainty. First, based on geodetic and seismic data we estimate the geometry of the involved faults assuming rectangular planar fault-surfaces with uniform slip. Then, we use only the geodetic data to estimate spatially variable final (static) slip on the geometry of inferred faults. Finally, we use the thus obtained static slip distribution to inform the Bayesian inference for the kinematic rupture evolution of the earthquake based on both the geodetic and seismic data, which we refer to as the finite-fault inference in the following.

Assuming a 1-D layered elastic half-space for the Earth structure (Khrepy et al., 2016), we calculate Green's functions (GFs) with 1 km spacing (Heimann et al., 2019) for the geodetic and the teleseismic data using the codes PSGRN/PSCMP (Wang et al., 2006) and QSSP (Wang et al., 2017), respectively. Sampling frequencies for the seismic GFs are 1 Hz and 4 Hz for the geometry and finite-fault inferences, respectively. To assess the fit of synthetic data $\mathbf{d}_{syn}$ to the observed dataset $\mathbf{d}_{obs}$ of waveforms and/or displacement maps we calculate the weighted variance-reduction (VR, Cohee and Beroza, 1994):

$$VR = \left(1 - \frac{(\mathbf{d}_{obs} - \mathbf{d}_{syn})\mathbf{C}^{-1}(\mathbf{d}_{obs} - \mathbf{d}_{syn})}{\mathbf{d}_{obs}\mathbf{C}^{-1}\mathbf{d}_{obs}}\right) * 100 \quad (1)$$

where $\mathbf{C}^{-1}$ is the inverse of the data covariance matrix. The closer the VR is to 100% the better the data are explained by the synthetics.

2.2.1 SAR data

We derived surface-displacement maps from five interferometric pairs of synthetic aperture radar (SAR) data and one SAR pixel offset map (Fig. 2) by using the GAMMA software (Wegmüller, 1998). Topographic phases have been evaluated and removed from the interferograms based on the SRTM digital elevation model (Farr et al., 2007). To increase the signal-to-noise ratio (SNR) the interferograms have been multilooked to ∼90 m x 90 m pixels and filtered with an adaptive phase filter (Goldstein and Werner, 1998). Finally, the filtered interferometric phases were unwrapped using a minimum cost flow algorithm (Chen and Zebker, 2001). To prepare the data for parameter inference, we reduced the number of pixels in the unwrapped interferograms by using the quadtree subsampling algorithm (Jónsson et al., 2002) and estimated the full data variance-covariance matrix following Sudhaus and Jónsson (2009) and Isken et al. (2017). The data from tracks 254 and 114 (Fig. 2) have not been used before in previous studies of this earthquake (Supplement Tab. S2). Note that due to irregular acquisition times (Supplement Tab. S1), the interferograms contain up to 24 and 21 months of pre- and post-seismic deformation, respectively. The influence of possible secondary sources of deformation on estimated parameters is discussed in sec. 4.3.

2.2.2 Teleseismic data

We used data from 27 broad-band seismic stations at teleseismic distances of $26.5° - 91.0°$ and $29.5° - 87.0°$ for the P and S wave data, respectively (Fig. S1,c & d). The data have been restituted to displacement and rotated to the radial, transverse and vertical (RTZ) source-receiver geometry. We applied band-pass filtering to contain waves with periods of 100s to 20s and 100s to 2s (i.e. band-pass filter between 0.01-0.05 Hz and 0.01-0.5 Hz) for the geometry and finite-fault inferences, respectively.

2.2.3 Inference of fault geometry

We estimated geometry and fault-slip kinematics of rectangular fault segments considering the following parameters for each fault segment: depth, length, width, slip, strike-, dip- and rake-angles. Assuming a half-cosinuidal source time function (STF, Lay et al., 2010) we also estimated nucleation-times and slip-duration of each fault segment. The top-center points of the fault segments have been fixed to be located on the mapped faults and surface ruptures of the earthquake (Ribot et al., 2021; Lefevre, 2018, i.e. faults annotated *AgF*, *AnF* and *mapped ruptures* in Fig. 1). In addition, we estimated hierarchical parameters for each interferogram, i.e. an offset and two ramp parameters (in the azimuth- and range directions) to mitigate effects of long wavelength atmospheric phase delays as well as inaccurate satellite orbit geometries. For the teleseismic data we estimated time shifts for each seismic station and waveform (P and S) to partially account for errors in the Green's Functions caused by lateral Earth-structure heterogeneities (Mustać et al., 2020; Vasyura-Bathke et al., 2021). We also estimated a noise-scaling parameter for each dataset residual to account for data and theory errors (Vasyura-Bathke et al., 2020, 2021). This setup yielded a total of 138 random variables, sampled from uniform distributions, to be constrained for the inference solution space which we explore using a sequential Monte-Carlo algorithm (Moral et al., 2006; Minson et al., 2013; Vasyura-Bathke et al., 2020).

2.2.4 Finite-fault inference

We employed the results from the geometry inference and fixed the fault segments geometry to the maximum a-posterior (MAP) solution. We then extended the fault segments in length and width in each direction, as fault dimensions are commonly underestimated applying the uniform-slip assumption, and discretized the fault segments with rectangular patches of 5.0 km. In total, we inferred >700 uniform distributed unknown parameters that comprise the slip in strike-parallel and down-dip directions, rupture duration, and rupture velocity on each patch. In addition, we estimated Laplacian smoothness regularization factors in conjunction with the location and time of separate rupture-nucleation points, i.e. one for each fault segment. We constrained the prior for the rupture-nucleation times for each fault segment based on a-priori information from the back-projection semblance maps. While we fixed the previously determined ramp parameters, we estimated noise-scaling factors for each dataset as well as the time shifts for each waveform applying sequential Monte Carlo sampling of the solution space.

3 Results

3.1 Backprojection source imaging

Our backprojection results (Fig. 3a) map several regions of coherent high-frequency seismic energy radiation moving from south to north along the Gulf of Aqaba. The initial high-frequency energy was coherently radiated at -8 to -4 s with respect to gCMT time 04:15:26 near the southern end of the Aragonese fault (AgF), followed by an apparent gap of seismic radiation with a duration of around 6 s (Fig. 3). Subsequently, high-frequency energy was radiated on the adjacent northern continuation of the AgF at around 2-4 s and on the east coast of the Gulf at 4-6 s, nearby the observed surface fractures reported by Lefevre (2018). Finally, seismic radiation was again likely generated on the AgF migrating northward between 6-12 s. The high-frequency energy radiation at 10-12 s is possibly an artefact of the processing and discussed in detail in sec. 4.2.2.

3.2 Geometry of fault segments

We defined two fault-geometry setups, each with a different number of fault segments, to investigate the impact of additional source complexity for explaining the data. The first fault geometry comprises two fault segments: a single long off-shore segment on the AgF and a short on-shore segment located on the eastern gulf coast where surface ruptures have been mapped Lefevre (2018) and where we found coherent high-frequency energy radiation in the semblance maps (Fig. 3). The second fault-geometry setup comprises three fault segments, two shorter off-shore segments (on AgF and AnF) and the segment on the eastern shore (Fig. 3). We then estimate the fault geometric parameters, strike-angle, dip-angle, width and length, while other inferred parameters, such as slip, nucleation-time and slip-duration were included to not bias the estimation of the fault geometric parameters. Their inferred posterior probability densities (PPD) thus provide initial estimates only, as they were refined in the following finite-fault inference.

3.2.1 Two fault segments

For the fault-model setup comprising two fault segments (Fig. 3a) the posterior ensembles show that the first segment (offshore, red) is well constrained. It dips $\sim$70°-72° towards west, and has length L of $\sim$50-53 km and width W >26 km (Supplement Fig. S2). The second segment on the eastern shore of the Gulf (blue) is poorly constrained and dips $\sim$80 to 90° towards west. The PPD for fault length L (between 5 - 8km) is constrained by the prior information of the surface structures at the lower end of the distribution (Supplement Fig. S3). Whereas, the fault widths of above 9-10 km are more likely. The fault strike-angle was constrained to be between 200-210° based on the a priori structural information, hence the PPD was truncated at $200°$. While the offshore segment shows well constrained strike-slip motion (rake-angle of $-7°$ to $-6°$) the eastern segment shows mostly dip-slip (rake-angle of $\leq -80°$). The weighted variance reductions (VRs; Eq. 1) of the geodetic data are $\sim$60$\%$ for most of the interferometric pairs (Supplement Figs. S4, S5), but lower for the amplitude offsets that are overall noisy. Highest correlated residuals are located on the western shore of the Gulf. In general, the VRs for seismic data are high for P and S phases, i.e. between $75 - 95\%$ (Supplement Figs. S6, S7). However, there are

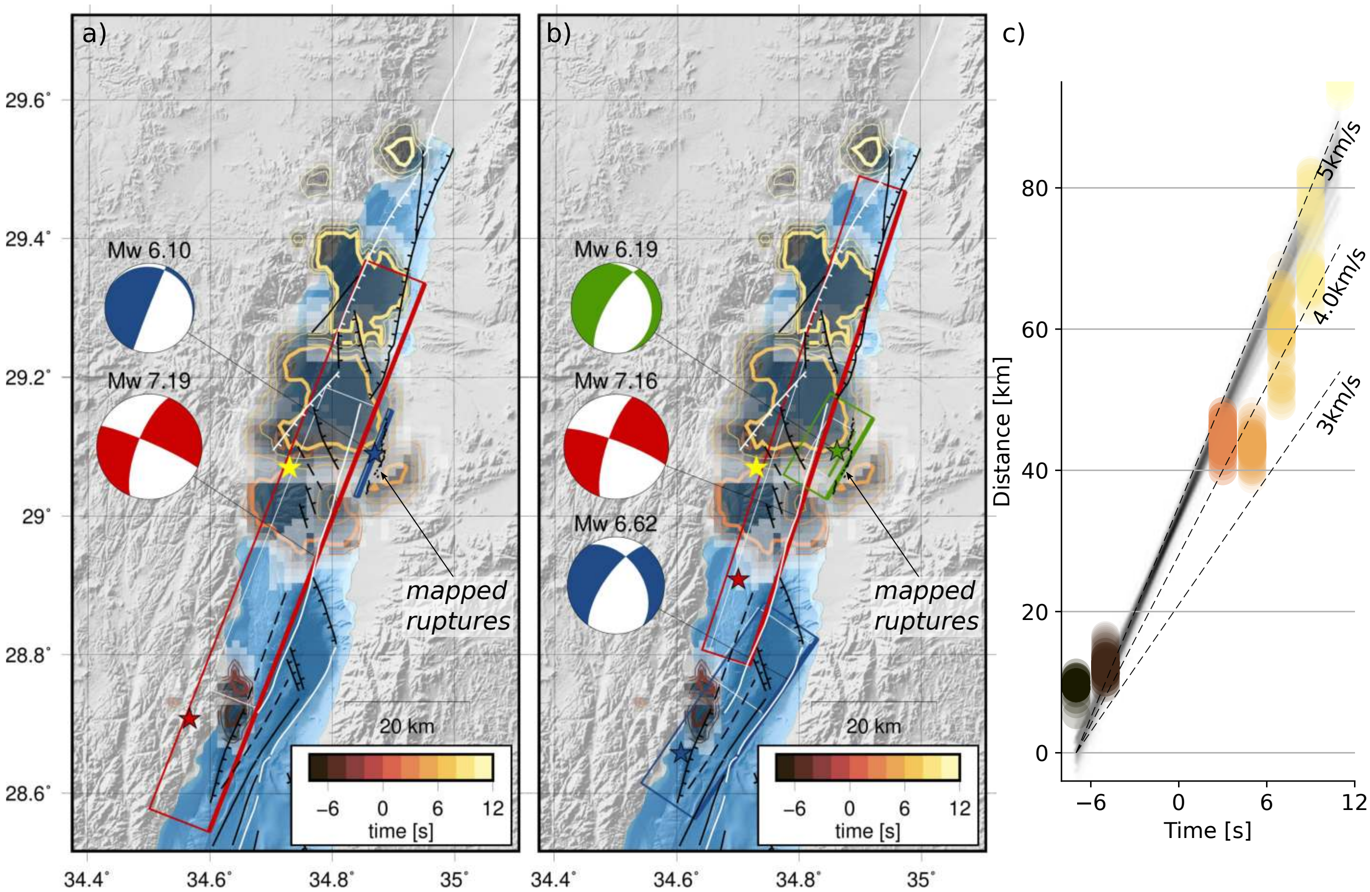


Figure 3 Results of the backprojection analysis showing the temporal evolution of multi-array semblance with respect to the global centroid moment tensor time 04:15:26. Thick colored contour lines show the 90% quantile of the ensemble of multi-array semblances (white-to-gray areas), whereas thin contour lines indicate the subsequent 60% and 30% quantiles. Coloring and annotation of fault structures is identical to Fig.1, slightly modified after Lefevre (2018) and Ribot et al. (2021). The colored rectangles are the a) two and b) three fault segments used for the estimation of the rupture evolution, expanded in size from the estimated faulty geometry (thin gray rectangles), and the colored stars show the respective estimated rupture nucleation points on each segment. c) Ensemble of multi-array semblances projected along the axis of the gulf (strike of $18°$ East of North) referenced with respect to the earliest semblance between -8 and -6 s. Dashed black lines indicate steady rupture velocities of 3, 4 and 5 km/s while shaded grey lines show averaged velocities for the whole ensemble of semblances.

a few stations that show remarkably low variance reductions, e.g. KMBO and BOSA (Fig. 4a).

3.2.2 Three fault segments

For the fault-geometry setup comprising three segments (Fig. 3b) the strike and dip angles of the northern segment (AgF) are well constrained at ∼198° and $75 - 76°$, respectively. The strike- and dip-angles of the southern segment (AnF) hits the chosen upper bound of the prior distribution at 215° and 70° constrained by geologic information (Ribot et al., 2021, Figs. S8, S9). While the northern segment shows predominantly strike-slip motion, the southern segment (AnF) has a larger normal component of slip (rake-angle of ∼-50° to -48°) compared to the northern segment (rake-angle of ∼-9°). The eastern on-shore segment (green) again shows predominantly normal slip (rake-angle of ∼-75°), Fig. S10). The variance-reduction (VR) values for three out of six geodetic datasets are higher by 5-12% (up to ∼76%) compared to the two-segment geometry (Supplement Figs. S11 and S12). For the seismic data the variance reductions are in general only slightly higher to those of the two-segment case. Stations KMBO.Z and BOSA.Z that showed low VRs for the two-segment setup have up to 70% higher variance reductions for the three-segment model. Moreover, other stations that show complexity in early P-phase arrivals (MSEY.Z, TATO.Z, HYB.Z) are explained significantly better in terms of amplitude and number of wave cycles, although in terms of VR these are only 5-7% higher (Fig. 4a). These improvements in VRs in comparison to the two-segments setup are significant and support the notion of geometrically complex faulting off-shore on the AnF and the AgF.

3.3 Finite-fault inference

3.3.1 Two fault segments

For the ensemble of two-segment finite-fault solutions, rupture initiates on the offshore segment at around -9.5 s to -8.2 s (wrt. gCMT time 04:15:26; Fig. 5a,b) between 25.5 km and 29.8 km depth, and then spreads unilaterally northeast across the segment. The onshore segment starts rupturing at ∼8-8.9 s, nucleating at shallow depth of around 1.0 km to 5.2 km and propagating

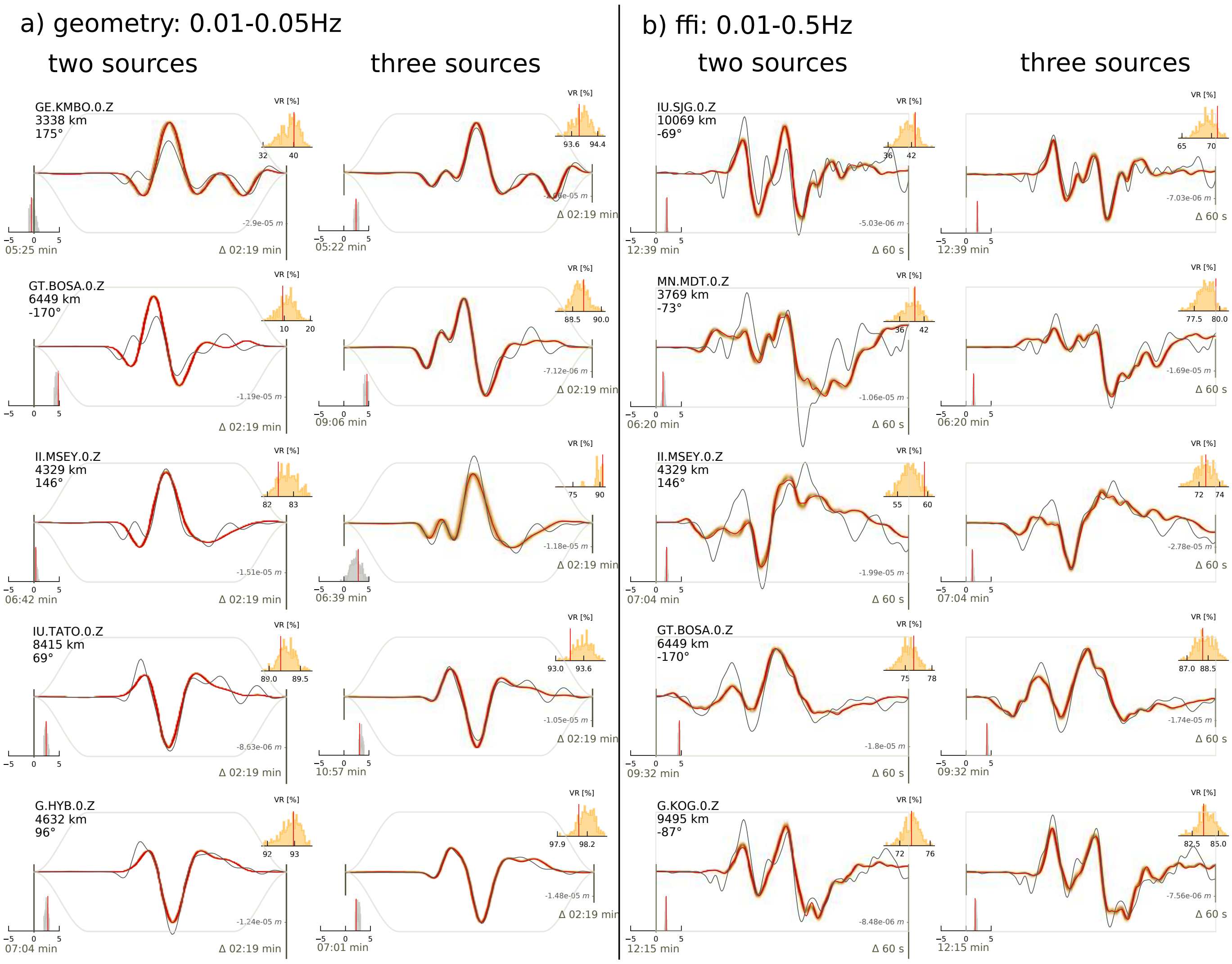


Figure 4 Waveform fits of several selected stations for the a) geometry and b) finite-fault inversion inference of the 1995 Gulf of Aqaba earthquake. The solid gray lines show the filtered data of the vertical component tapered around the P-wave arrival. The red continuous lines show the synthetic waveforms derived from the MAP solution, whereas, brownish shaded colors indicate the synthetic waveforms derived from the full posterior ensemble of parameters.

unilaterally across the fault. The highest values of fault slip (> 1.8 m) occur from the surface to a depth of ~9 km on the central part of the offshore segment along a distance of ~25 km. While fault motion is mostly strike-slip, some shallow slip with significant normal component is found. The timing and thus the velocity of the inferred rupture front towards north is consistent with the semblance maps from the BP, i.e. the rupture front reaches the region with high slip amplitude at ~4 s. The moment-rate function reveals that seismic-moment release started gradually and linearly (Fig. 5a) and reached its maximum with several sharp peaks between ~6-8 s. This is followed by a fast decay of moment release until ~24-25 s. The geodetic data have an average VR of $49.9 - 52.3\%$ (Supplement Figs. S15, S16), where the amplitude offsets with the highest noise show the lowest VR. Most noticeable residuals are again located on the western shore of the Gulf. Seismic data are well explained with an average VR of $74.2 - 77.8\%$ (Supplement Figs. S17, S18). Here, both P and S phases are well modelled in general, but amplitudes of the main pulse and early P-phases are often biased resulting in low VRs of $36 - 60\%$, e.g. SJG.Z, MDT.Z, MSEY.Z (Fig. 4b).

3.3.2 Three fault segments

In the ensemble of three-segment finite-fault solutions, rupture initiates on the southern segment at around -8.2 s to -7.8 s (wrt. gCMT time 04:15:26; Figs. 6a,b, 7) between 18.8 km and 22.5 km depth, and spreads unilaterally towards the Northeast across the segment. The second rupture then nucleates on the northern segment at ~-1.5 s to -1.1 s between 19.3 km and 23.2 km depth and propagates in rather unilateral direction. Finally, the third, eastern segment starts rupturing between 3.7 s and 3.9 s at 4.7-9.4 km depth with unilateral rupture towards the down-dip direction (see also video in supplement). The slip is dominantly left-lateral on the southern and northern segments, but there is a notable normal component on all three segments (Fig. 6b). While the normal component is large at greater depths on the southern segment, it is high at shallow depths on the northern and eastern segments, respectively. The region with largest slip of >1.8 m, a length of ~25 km

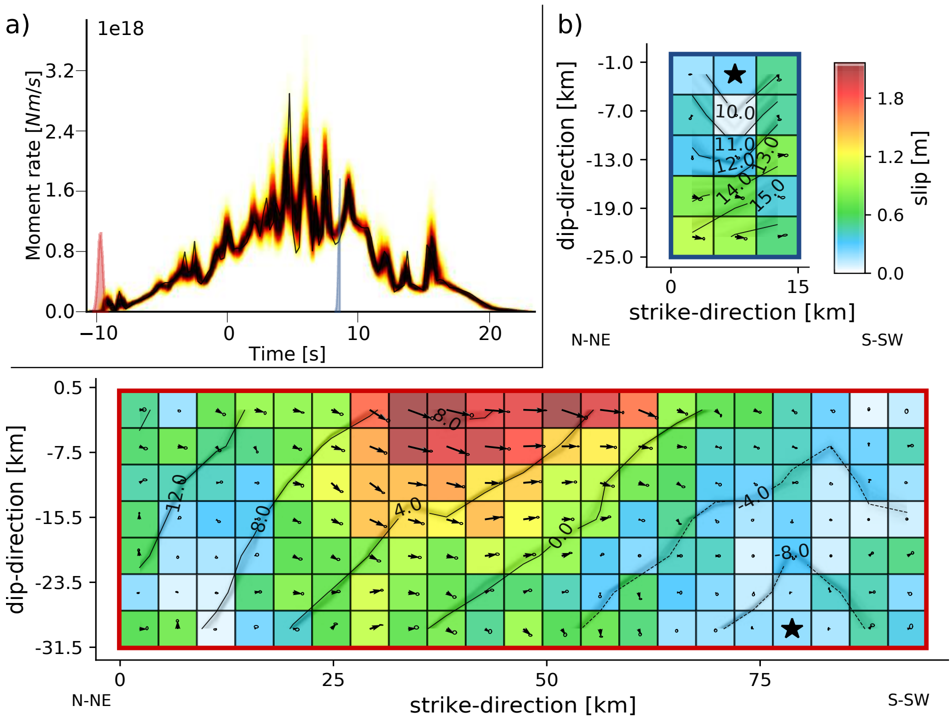


Figure 5 Estimated spatiotemporal rupture evolution for the case with two fault segments. a) Ensemble of moment-rate functions with darker colors highlighting high-probability moment-rates. The maximum a-posterior moment rate is shown by the solid black line. Colored histograms (next to the moment-rate) show estimated rupture nucleation times of each fault-segment. b) Kinematic slip-distribution of the 1995 Gulf of Aqaba earthquake. Patch colors and arrows mark the MAP solution, whereas black ellipses mark the 95% standard deviation of slip. Black stars indicate the inferred rupture nucleation point on each fault segment and the black continuous lines show the MAP rupture front with the annotated timing. Uncertainty in rupture-front propagation is shown by fuzzy light-gray isolines indicating lower probability. All times are with respect to the global centroid moment tensor time 04:15:26.

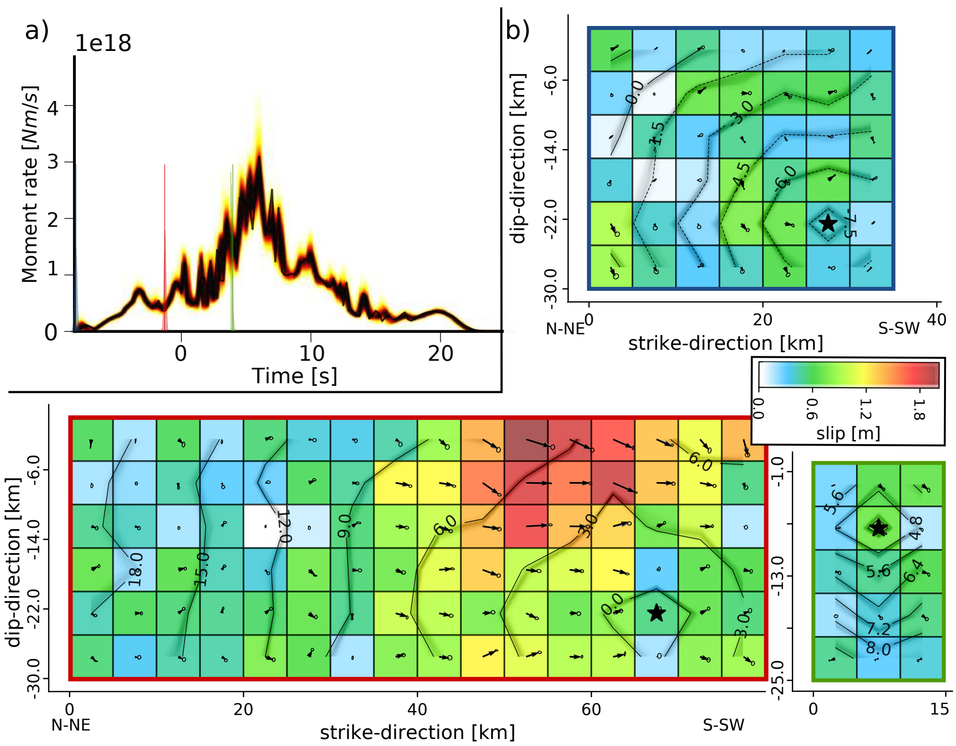


Figure 6 Same as Fig. 5, except for the case of three fault segments.

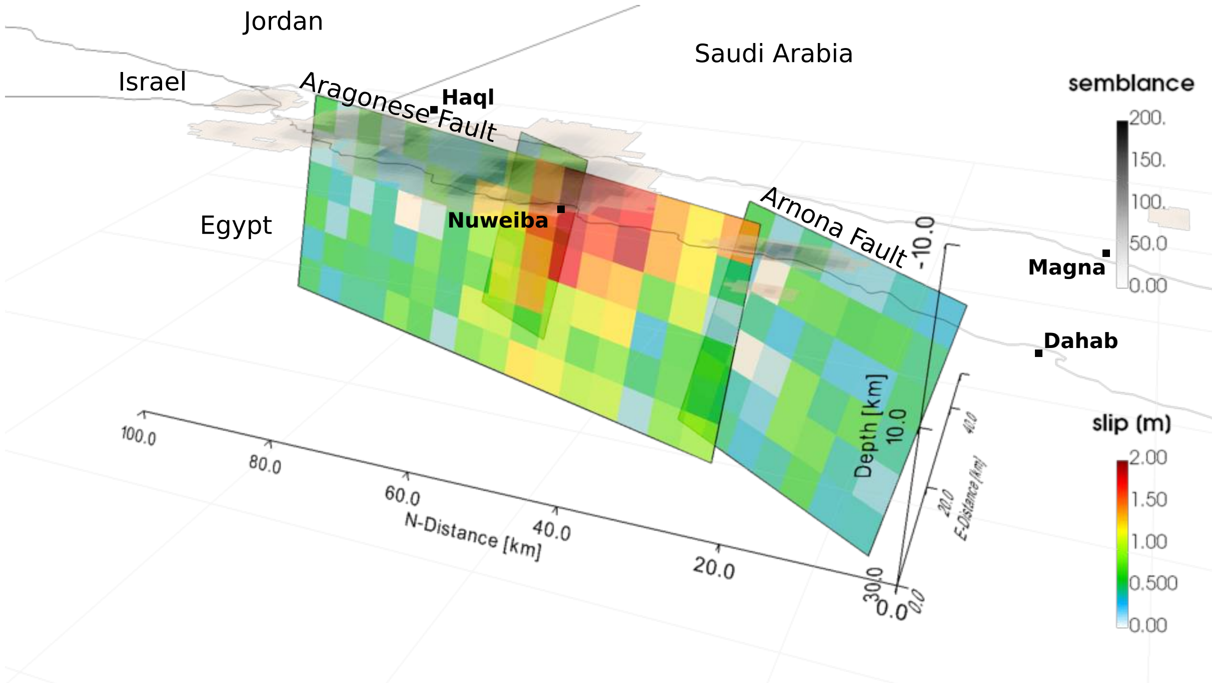


Figure 7 Estimated slip-amplitude distribution for the three fault-segment case in 3D perspective view. Greyish colors show the back-projection semblance maps from Fig. 3. Thin black lines show coastlines and country borders. A video of the temporal rupture evolution can be found in the supplement.

and is located near the southern end of the northern segment, extending from the surface downwards to ~9-15 km depth. On the southern segment, maximum slip is ~1 m close to the hypocentre at depths between 15 km and 22.5 km. The moment release started gradually and increased rapidly once the northern segment started rupturing (Fig. 6a). Most of the moment had been released at ~20s. The rupture velocity is noticeably faster on the southern segment ~3.2-4.7 km/s compared to that on the northern segment ~2.4-3.5 km/s. It is slowest on the eastern segment ~2.3-3.7 km/s (Figs. 6b, S19). In general, rupture velocity is better constrained close to the rupture nucleation points compared to further away.

Overall the geodetic data are slightly better explained by the three-segment model rather than by the two-segment setup with an average VR between $53.7-55.9\%$. The largest residuals are located on the western shore of the Gulf (Supplement Figs. S20, S21). Seismic data are similar or slightly better explained than for the two-segment model with an average VR between $77.7-80.1\%$ (Supplement Figs. S22, S23). Most noticeable improvements (change in VRs between ~10-40%) to the two-segment setup are for P-phases in the amplitudes of early arrivals e.g. SJG.Z, BOSA.Z, KOG.Z (Supplement Figs. S22) as well as for the main pulse, e.g. MDT.Z, MSEY.Z.

4 Discussion

We derived kinematic finite-fault rupture models for the 1995 Gulf of Aqaba earthquake that were estimated, to our knowledge, for the first time by the joint use of geodetic and seismic data through Bayesian inference. Taking into account only the ability of the presented model(s) to explain the data in terms of variance reduction allows to reach a conclusive answer on the suitability of either the two or the three fault segment setups. While the three-segment setup only slightly better explains the geodetic data, it significantly better explains the seismic data. Calculating the Bayesian Information Criterion (BIC, Schwarz, 1978) for both models revealed that the three-segment configuration with 718 unknown parameters (153 patches) had a lower value than the two-segment configuration with 762 unknown parameters (162 patches). This result supports the use of the three fault segment configuration. Still, there are features in the two-segment geometry setup that are preferable to the other and vice versa. In the following section we discuss these features in detail.

4.1 Fault geometry

We find that geometric fault complexity approximated by the three-segment model is needed to explain especially early seismic phases. These indicate rupture on the northern end of the AnF, dipping westward towards the Aragonese Deep. For the gulf our inferred fault geometry shows a westward inclined segment along the entire fault length, although it was mapped in the south to be dipping eastward towards the Aragonese Deep and to be dipping westward towards the Elat Deep in the north (Ribot et al., 2021). Through forward modeling we tested the possibility for a vertical or eastward inclined southern end of the AgF segment (Fig. S24). It is highly unlikely that an eastward inclined fault has been activated during the earthquake as it would cause a clear data misfit on the eastern coast of the Gulf, assuming a fault segment dipping with $80°$ towards the East. Assuming a vertical fault segment does not cause as large a misfit to the geodetic data (Fig. S24a,b), but the seismic data (Fig. S24c) is poorly explained on the Z-components. In conclusion, we propose that the Aragonese fault is curved in along-strike direction and that it is dipping westwards in the north and that it is subvertical close to the Aragonese Deep in the south. Depending on the amount of fault curvature on the inferred fault segments (which we assumed as planar in our models), the inferred distribution of slip may therefore be biased (Dutta et al., 2021).

4.2 Temporal rupture evolution

4.2.1 Artefacts in Finite Fault Inference

According to our source-inference results, several unilateral rupture fronts on the AnF and AgF, illustrated by the three-segment fault geometry, better explain the seismic data than a single unilateral rupture front involving only the AgF. However, there are second order features in the inferred kinematic finite fault setups that could rather be attributed to artifacts, potentially caused by theory error. Firstly, the deep rupture nucleation on the lower edges of the offshore faults in both fault-geometry models at depth $>\sim$22-25 km is rather unrealistic as the crustal thickness in the Gulf is reported to be $\sim$20 km. Secondly, the moment rate functions (Figs. 5a, 6a) show a long tail of moment release up to a total rupture duration of $\sim$30 s. This may be at least partially a result of the assumed sinusoidal functional form of the local source-time function on each fault patch (Meier et al., 2017). An alternative source-time function more consistent with earthquake rupture dynamics, e.g. the regularized Yoffee function (Tinti et al., 2005), may result in a shorter estimated total moment-rate function which would be more consistent with the apparent rupture duration of $\sim$20 s imaged by BP. Thirdly, the rupture velocity on the onshore segment for the two-segment setup is very slow $\sim$2 km/s and consequently, the rupture duration of that fault segment is long compared to its size (Kanamori and Brodsky, 2004). However, rupture-velocity and duration are source parameters which are potentially biased or influenced due to over-fitting to compensate some of the theory error. To improve this, the kinematic evolution of the rupture could be better resolved by utilising regional seismic data, which would likely reduce bias and artifacts in the estimated parameters.

4.2.2 Implications from back-projection

The employed BP method maps the coherent seismic radiation of P-wave energy related to changes in fault geometry and changes in rupture velocity, rather than the amount of fault-slip. Therefore, BP semblance is indicative of rupture nucleation, rupture arrest and kinks or bends in the fault geometry, however, the semblance values are not directly proportional to radiated seismic energy. The BP semblance map indicates a complex change of coherent energy radiation during the rupture process, and to first order resembles the inferred finite-fault models. Rupture velocities obtained from the BP results agree with those inferred in the finite-fault inversions: 3.3-3.7 km/s around the southern fault segment, 3.5-4.2 km/s around the northern fault segment and 2.8-3.3 km/s on the eastern fault segment.

One apparent discrepancy in the BP results compared to the inferred rupture models is the last mapped semblance at 18-20 s after rupture onset at the northern end of the AgF, onshore. These semblances may indicate that some fault segments in the north of the Gulf, such as the EF, slipped during the earthquake or that rupture was shallower or deeper than the depth of the grid used for the BP. No studies have reported a fault segment rupture this far north on the EF. Therefore, a mismapping of this semblance seems likely. Previous studies have shown that the choice of grid depth and the deviation of the source depth from the grid depth have a strong effect on the location of the mapped semblance (Steinberg et al., 2022; Daout et al., 2020), so this seems to be the likely cause. However, waveform coda and depth phases can also contaminate the BP results, especially in the later stages of the rupture process. In particular, depth phases may have relatively large amplitudes compared to direct phases, yet, for shallow earthquakes these phases arrive close in time in the seismic records. This results in unwanted side lobes either in parallel or at an acute angle to the fault depending on the station-array geometries. This effect can only be suppressed by using multi-phase semblance (Steinberg et al., 2022), which is hindered for the 1995 Gulf of Aqaba earthquake by limited data availability. Furthermore, the depth phase separation is challenging, particularly for strike-slip earthquakes that generate relatively large sP phases. Therefore, caution is necessary when interpreting the semblance maps of the later time steps in the BP.

To compare the two inferred kinematic finite-fault models to the BP results we carry out synthetic BPs on synthetic waveforms calculated by using the model parameters of the two- and three-segment finite-fault models (Fig. 8a, b). We used the same station setup and data processing as for the real data BP. Synthetic waveforms were calculated at a sample rate of 4 Hz using a Green's function store calculated with QSEIS and the AK-135 Earth structure model (Kennett and Engdahl, 1991). For both synthetic BPs we observe early bilateral rupture due to the abrupt stopping of the rupture at the model-fault edges. The observed semblance map of the real data BP at the northern end of the Gulf was not reproduced by the synthetic BP of either finite fault model.

The simple kinematic model used to calculate the synthetics produces sharper start and stop phases than actually observed (Steinberg et al., 2022), as the modelled fault ends abruptly, whereas in nature there would be tapering. Furthermore, the synthetic semblance maps produce much sharper semblance patches in comparison to the real data backprojection. The reason for this is due to the lack of data noise in the synthetic backprojection and, the fact that semblance is mapped onto a single fixed depth, which causes a blurred semblance in the real data backprojection.

In the synthetic BP of the two-segment finite-fault model the apparent rupture velocities are slow compared to the real data BP. The synthetic BP of the three-

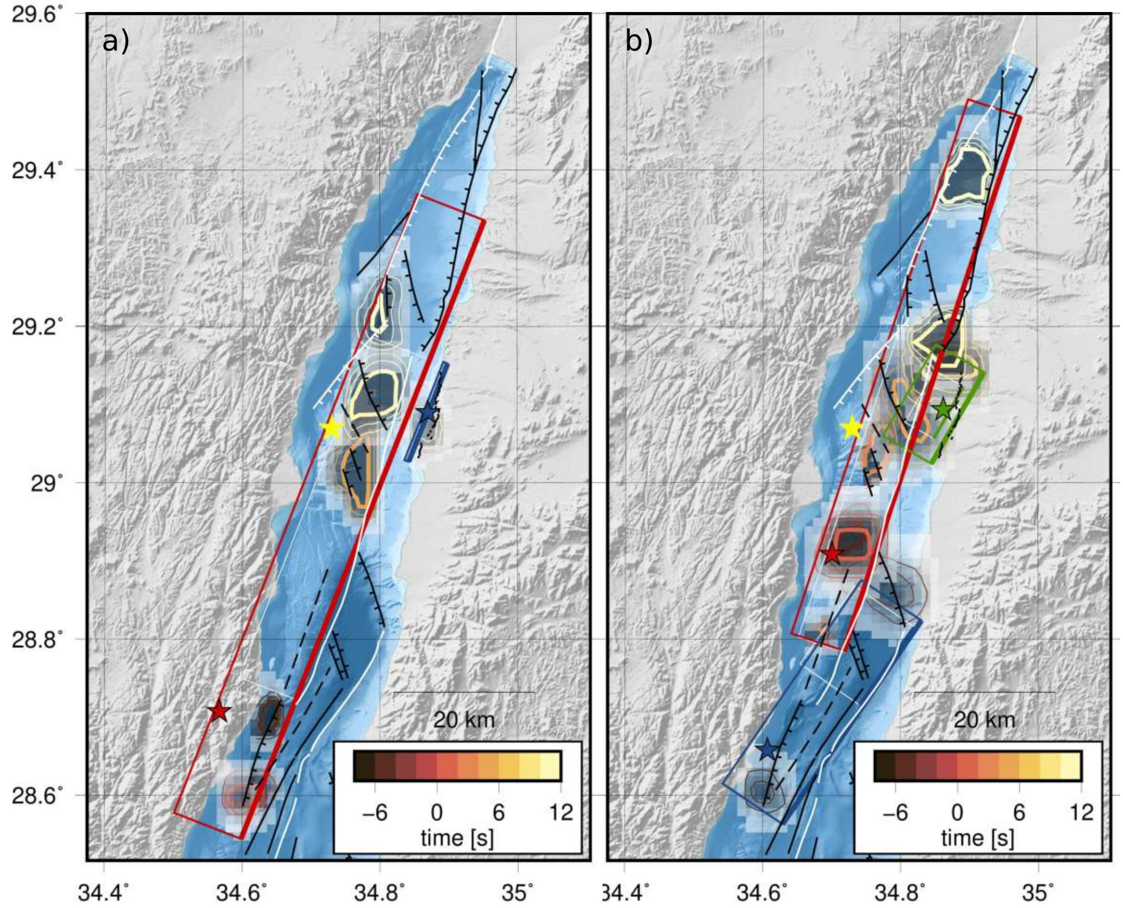


Figure 8 Results of backprojection analysis of synthetic waveform data generated based on the MAP model of the a) two-segment and b) three-segment setups showing the temporal evolution of multi-array semblance with respect to the global centroid moment tensor time 04:15:26. Details are similar to Fig. 3

segment finite fault model (Fig. 8b) is in better agreement with the observed BP results (Fig. 3b), especially for the complicated semblance distribution in the center portion of the AgF. These probably indicate a change in rupture speed or fault geometry, e.g. in the form of a fault bend. The synthetic BP of the three fault segments also agrees well with observed complexity of semblance mapping in the area of the eastern fault segment.

4.2.3 Rupture nucleation and fault segmentation

Our scenario with multiple rupture episodes, such as in the three-segment setup where rupture began on the AnF and subsequently propagated to the AgF, shows that rupture speed on the AnF is too slow for nucleation on the AgF to be triggered by unilateral rupture on the AnF or by the arrival of S-waves. The distance between the nucleation point on the AgF segment, which has a nucleation time of ∼1.5 s and the rupture front on the AnF segment at the same time, we estimate a rupture jumping distance between ∼11-37 km. This is on the order of the 20 km jumping distance postulated for the 2016 Mw 7.8 Kaikoura earthquake (Cesca et al., 2017; Shi et al., 2017) and larger than previously proposed maximum jumping distances of 3-4 km (Wesnousky, 2006). However, the rupture nucleation on the AgF could have been caused by dynamic triggering of P-waves emitted during the early stages of the AnF segment rupture. The P-wave velocity is between ∼6.5 and 6.9 km/s (Khrepy et al., 2016), such that a theoretical travel-time of ∼3.5-5.5 s would be required between nucleation points (Fig. 9; distance of 27-32 km). The difference between inferred nucleation times in our three-segment setup as well as the temporal difference between the BP semblance maps are ∼6-7 s (Figs. 3-5), and thus they are within a plausible range for this required travel time. Using QSEIS we simulated the seismic wavefield generated by slip on all patches of the first fault-segment and then received at the second nucleation point. This allows for estimating the dynamic Coulomb stress imposed at the nucleation point on the second fault segment (Fig. 9). The second rupture on the northern segment initiated at a Coulomb stress change of ~ 0.3 MPa which is consistent with previously reported dynamic Coulomb stress changes needed to trigger earthquake rupture (e.g. Antonioli et al., 2006). Thus, it seems plausible that rupture on the AgF segment was triggered by the early pulses of deformation associated with the seismic phases that have been emitted from the AnF segment. Another possibility is that a small, unresolved normal-fault segment acted as an intermediary and facilitated an apparent jump of the rupture. However, this would require that the rupture process on this small normal-fault segment was very smooth and did not emit any start or stop phases that could be resolved through backprojection using the available data.

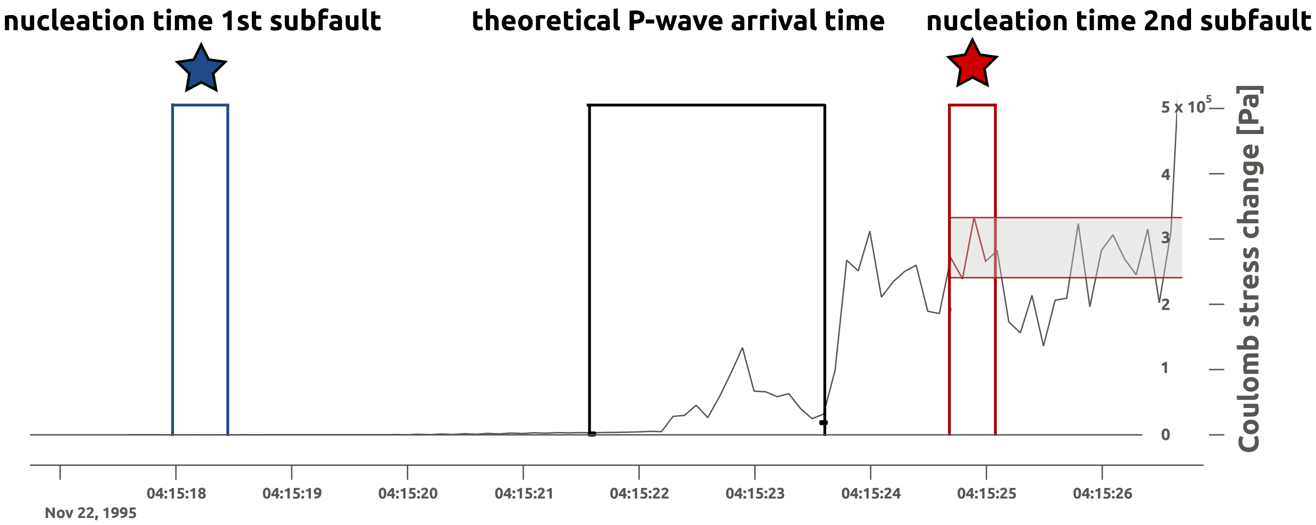


Figure 9 Time-dependent Coulomb stress changes at the second nucleation point of the three-segment fault geometry imposed by all source patches from the southern fault segment. The blue and red boxes mark the uncertainty of the inferred nucleation time of the southern and northern fault segments, respectively. The area indicated by the black box marks the window of theoretical P-wave arrival times (for the posterior ensemble of models) for a P-wave traveling from the hypocenter to the second nucleation point on the northern fault-segment. The shaded grey box marks the interval of inferred Coulomb stress-change during the initiation time of the northern fault segment.

4.3 Secondary sources of deformation

Residual displacements of pixel offsets and interferograms are in general relatively high (Supplement Figs. S11, S12), especially on the western coast of the Gulf. On the other hand, residuals of seismic waveforms are low (Figs. S17, S18, S22, S23). Interferometric pairs cover significant time-spans where the secondary acquisitions have been acquired 5-6 months after the mainshock (Supplement Tab. S1). During this time, post-seismic deformation processes have occurred that are not included in the presented co-seismic kinematic finite-fault models. Especially, four shallow normal faulting aftershocks with magnitudes $M_w \geq 3.9$ (Baer et al., 2008, Supplement Tab. S2) have been reported (Hofstetter et al., 2003) on the western coast of the Gulf where the large residual deformation is present in the SAR data (Figs.1, S20, S21). In conjunction with post-seismic deformation such as afterslip (Baer et al., 2008) these can account for ~5-7 cm of displacement, which is the bulk part of residual displacements in the interferograms. Thus, the joint inference of seismic and geodetic data played an important role for disseminating the contribution of these signals.

4.4 Ground-motion map from finite fault inference

The surface effects of an earthquake on infrastructure and population are typically evaluated using ground motion maps. The inferences from our finite fault modelling of the 1995 Aqaba earthquake can be used to produce an informed estimate of realistic peak-ground velocity (PGV) predictions for this particular earthquake. These may then serve as a basis for scenario calculations for seismic hazard assessment in the region.

We calculated a deterministic physics-based ground-motion map by simulating the seismic wavefield based on the spatiotemporal rupture evolution of the three-segment fault model (Fig. 6, Dahm et al., 2018). For the wavefield simulation we calculate Green's functions with QSEIS (Wang, 1999) with a sample rate of 20 Hz at a dense set of virtual receivers (1 km spacing) assuming the 1D-layered elastic Earth structure model used also in our FFI (Khrepy et al., 2016). We considered site effects through the shear-wave velocity in the thirty meters below Earth surface, Vs_{30}, derived from the topographic slope as a proxy (Wald and Allen, 2007). We calculate the amplification ratio A_{30} between the predicted Vs_{30} and the shear-wave velocity of the uppermost layer resulting from our Green's functions. After bandpass filtering the simulated waveforms between 1-8 Hz and subsequent rotation to the RTZ coordinate system the expected PGV at each grid point is calculated from the maximum of the geometric mean of the absolute horizontal components R and T (Wald and Allen, 2007). The predicted PGV's (Fig.10) close to the source region reach 2 m/s. Adjacent areas at intermediate distances of ~20 km perpendicular to the source region of the main active fault (red), which has the highest inferred amount of slip, still have predicted PGV values of up to 0.5 m/s. The rupture directivity of the Gulf of Aqaba earthquake was toward north, and hence the shaking in the northern gulf is stronger than in the southern gulf and adjacent coastal regions.

4.5 Comparison to published models

Earlier studies (Pinar and Türkelli, 1997; Klinger et al., 1999; Hofstetter et al., 2003; Shamir et al., 2003; Baer et al., 2008) modelled the Aragonese fault as a single planar fault, similar to the offshore segment of our two-segment fault geometry setup. In published finite-fault

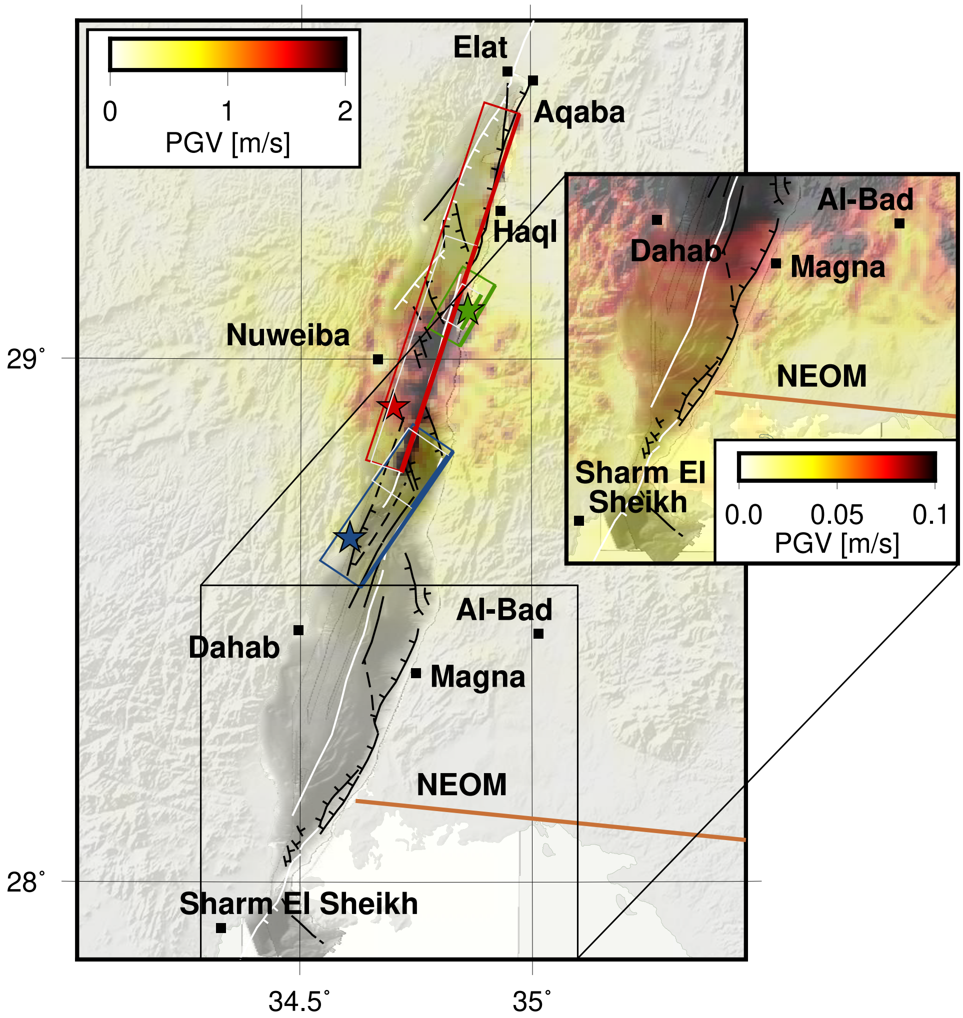


Figure 10 Map of estimated peak ground velocity (PGV) from seismic wavefield simulation (frequency range 1-8 Hz) for our three-segment rupture model of the 1995 Gulf of Aqaba earthquake. The colored rectangles mark the three fault segments. The black rectangle shows the zoom-in around the area of the urban project NEOM; note the different scale of colormap in the zoom-in. "The Line" (brown) is a $\sim 170\,km$ long city under construction. Coloring and annotation of fault structures is identical to Fig.1.

slip models the large-slip area is located at depth between 5 and 15 km and peak slip-amplitudes are larger by ~1 m (Hofstetter et al., 2003; Baer et al., 2008) compared to our result (Fig. 5b, 6b). In contrast to published models our slip models show a notable component of normal slip at shallow depths to ~8 km depth. While our findings are in line with most of the previous studies that the earthquake consisted of several subevents (Fig. 1, Supplement Tab. S2), they differ in the locations of the sub-events (Pinar and Türkelli, 1997; Klinger et al., 1999). While our three-segment model supports the proposed hypothesis that rupture initiated on the AnF in the south and then jumped to the AgF it does not support continued jumping to the EF in the north. Our results support that the deformation on the western coast of the Gulf was largely due to post-seismic deformation and aftershocks (Baer et al., 2008) rather than co-seismic mainshock slip. This post-seismic activity was predominantly characterised by vertical surface displacements, i.e. subsidence 20 km north of Nuweiba (Fig. S25). Furthermore, we propose that a fault segment on the eastern shore of the Gulf was active during the mainshock. Although, the inference of teleseismic and geodetic data do not allow to constrain this segment well, the BP results and the mapped structures (Fig. 1) indicate active faulting there. All these findings, i.e. shallow normal slip, normal faulting sub-event on the eastern shore and post-seismic normal faulting on the western shore are indicative of active tectonic extension during and after the Gulf of Aqaba earthquake.

While the inferred moment-rate function of our two-segment setup is similar to the one obtained by Hofstetter et al. (2003), i.e. roughly triangular symmetric with a rupture duration of ~25 s; the inferred moment-rate function for our three-segment setup is more similar to that of Pinar and Türkelli (1997). The total estimated moment magnitudes of our two and three-segment models are 7.24 and 7.27, respectively, which is slightly greater than previous estimates of 7.04-7.21 (Tab. S1).

5 Conclusions

We imaged the rupture of the 1995 Gulf of Aqaba earthquake using teleseismic multi-array backprojection. Mapped fault structures were used as prior information to constrain the location of activated faults in the gulf. We also estimated the kinematic finite-fault rupture evolution of the Aqaba 1995 earthquake using geodetic and teleseismic data jointly within a Bayesian inference process. We find that most of the rupture has occurred on the offshore west-dipping Aragonese fault which is curved in the along-strike direction. However, rupture initiated on the west-dipping Arnona fault south of the Aragonese fault. The temporal rupture evolution is complex and the inversion results support unilateral rupture originating close to the northern end of the Arnona fault and jumping over to the Aragonese fault that then again ruptured unilaterally towards the north. The backprojection results support this case. The earlier rupture in the south could have dynamically triggered the rupture on the Aragonese fault. In contrast to earlier studies we argue that a small onshore fault segment on the eastern coast of the gulf was seismically active during the event. While the event was predominantly strike-slip our models show that a significant portion of normal slip must have occurred along the Arnona and Aragonese faults. In conjunction with the sub-event on the eastern shore and post-seismic normal faulting on the western shore, our results suggest active tectonic extension of the gulf. Overall, this study presents new earthquake rupture models that describe the temporal rupture evolution of the 1995 Gulf of Aqaba earthquake. Especially, the postulated dynamic triggering between the fault segments should be taken into account for hazard models in the area, as this shows that seismic moment can be released faster over a shorter amount of time than during a purely unilateral rupture. Given large infrastructure projects such as NEOM that are actively being developed in vicinity of the Gulf of Aqaba with its geometrically complicated active fault system, the modelling of earthquakes in the region is important for updated and informed hazard assessment and for establishing scenarios of potential future earthquakes.

6 Acknowledgements

The authors thank František Gallovič, an anonymous reviewer and the editors for their comments which helped to improve the quality of the manuscript. HVB thanks Nicolas Castro-Perdomo for providing shapefiles of structures around the Dead Sea fault. This research was supported by King Abdullah University of Science and Technology (KAUST), under award numbers BAS/1/1353-01-01, BAS/1/1339-01-01, and OSR-CRG2020-4335. H.V-B was partially supported by Geo.X, the Research Network for Geosciences in Berlin and Potsdam under the project number SO_087_GeoX. A.S. was partially supported by the German Research Foundation DFG through an Emmy-Noether Young-Researcher-Grant (Number 276464525, "Bridging Geodesy and Seismology for improved and automated estimation of faulting events").

Data and code availability

The facilities of IRIS Data Services, and specifically the IRIS Data Management Center, were used for access to waveforms, related metadata, and/or derived products used in this study under https://ds.iris.edu. IRIS Data Services are funded through the Seismological Facilities for the Advancement of Geoscience (SAGE) Award of the National Science Foundation under Cooperative Support Agreement EAR-1851048. The post-processed InSAR data are available for download at https://zenodo.org/records/10462416 (Vasyura-Bathke et al., 2024). Bayesian inferences and multi-array teleseismic backprojection were performed using the Bayesian Earthquake Analysis Tool (BEAT; https://github.com/hvasbath/beat, Vasyura-Bathke et al., 2019, 2020) and Palantiri (https://github.com/braunfuss/Palantiri) (Steinberg, 2021; Steinberg et al., 2022), respectively. This work employed the open source libraries: *Numpy* (Harris et al., 2020), *Scipy* (Virtanen et al., 2020), *pyrocko* (www.pyrocko.org) (Heimann et al., 2017). Plots have been produced by using *Matplotlib* (Hunter, 2007) and the Generic Mapping Tools (GMT) (Wessel et al., 2013).

References

Abdel Fattah, A. K., Hussein, H. M., Ibrahim, E. M., and Abu El Atta, A. S. Fault plane solutions of the 1993 and 1995 Gulf of Aqaba earthquakes and their tectonic implications, 1997. doi: 10.4401/ag-3831.

Antonioli, A., Belardinelli, M. E., Bizzarri, A., and Vogfjord, K. S. Evidence of instantaneous dynamic triggering during the seismic sequence of year 2000 in south Iceland. *J. Geophys. Res. Solid Earth*, 111(3):1–14, 2006. doi: 10.1029/2005JB003935.

ArRajehi, A., McClusky, S., Reilinger, R., Daoud, M., Alchalbi, A., Ergintav, S., Gomez, F., Sholan, J., Bou-Rabee, F., Ogubazghi, G., Haileab, B., Fisseha, S., Asfaw, L., Mahmoud, S., Rayan, A., Bendik, R., and Kogan, L. Geodetic constraints on present-day motion of the Arabian Plate: Implications for Red Sea and Gulf of Aden rifting. *Tectonics*, 29:TC3011, 2010. doi: 10.1029/2009TC002482.

Baer, G., Shamir, G., Sandwell, D., and Bock, Y. Crustal deformation during 6 years spanning the M w = 7.2 1995 Nuweiba earthquake, analyzed by Interferometric Synthetic Aperture Radar. *Israel Journal of Earth Sciences*, 50(1):9–22, 2001. doi: 10.1560/12MA-85HY-WHXL-0GKA.

Baer, G., Funning, G. J., Shamir, G., and Wright, T. J. The 1995 November 22, M w 7.2 Gulf of Elat earthquake cycle revisited. *Geophys, J. Int.*, 175(3):1040–1054, dec 2008. doi: 10.1111/j.1365-246X.2008.03901.x.

Ben-Avraham, Z. Structural framework of the Gulf of Elat (Aqaba), Northern Red Sea. *J. Geophys. Res.*, 90:703–726, 1985. doi: 10.1029/JB090iB01p00703.

Castro-Perdomo, N., Viltres, R., Masson, F., Klinger, Y., Liu, S., Dhahry, M., Ulrich, P., Bernard, J.-D., Matrau, R., Alothman, A., Zahran, H., Reilinger, R., Mai, P. M., and Jónsson, S. Interseismic deformation in the Gulf of Aqaba from GPS measurements. *Geophys. J. Int.*, 228:477–492, 2022. doi: 10.1093/gji/ggab353.

Cesca, S., Zhang, Y., Mouslopoulou, V., Wang, R., Saul, J., Savage, M., Heimann, S., Kufner, S. K., Oncken, O., and Dahm, T. Complex rupture process of the Mw 7.8, 2016, Kaikoura earthquake, New Zealand, and its aftershock sequence. *Earth Planet. Sci. Lett.*, 478:110–120, 2017. doi: 10.1016/j.epsl.2017.08.024.

Chen, C. W. and Zebker, H. A. Two-dimensional phase unwrapping with use of statistical models for cost functions in nonlinear optimization. *J. Opt. Soc. Am. A*, 18:338–351, 2001. doi: 10.1364/JOSAA.18.000338.

Cohee, B. P. and Beroza, G. C. Slip distribution of the 1992 Landers earthquake and its implications for earthquake source mechanics. *Bulletin of the Seismological Society of America*, 84(3): 692–712, 06 1994. doi: 10.1785/BSSA0840030692.

Dahm, T., Heimann, S., Funke, S., Wendt, S., Rappsilber, I., Bindi, D., Plenefisch, T., and Cotton, F. Seismicity in the block mountains between Halle and Leipzig, Central Germany: centroid moment tensors, ground motion simulation, and felt intensities of two M3 earthquakes in 2015 and 2017. *J. Seismol.*, 22(4): 985–1003, 2018. doi: 10.1007/s10950-018-9746-9.

Daout, S., Steinberg, A., Isken, M. P., Heimann, S., and Sudhaus, H. Illuminating the Spatio-Temporal Evolution of the 2008–2009 Qaidam Earthquake Sequence with the Joint Use of Insar Time Series and Teleseismic Data. *Remote Sensing*, 12(17):2850, 2020. doi: 10.3390/rs12172850.

Dutta, R., Jónsson, S., and Vasyura-Bathke, H. Simultaneous Bayesian Estimation of Non-Planar Fault Geometry and Spatially-Variable Slip. *J. Geophys. Res. Solid Earth*, 126(7):1–28, 2021. doi: 10.1029/2020JB020441.

Fan, W. and Shearer, P. M. Investigation of backprojection uncertainties with M6 earthquakes. *J. Geophys. Res. Solid Earth*, 122 (10):7966–7986, 2017. doi: 10.1002/2017JB014495.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., and Alsdorf, D. The shuttle radar topography mission. *Rev. Geophys.*, 45(2005):1–33, 2007. doi: 10.1029/2005RG000183.

Goldstein, R. M. and Werner, C. L. Radar interferogram filtering for geophysical applications. *Geophys. Res. Lett.*, 25(21): 4035–4038, nov 1998. doi: 10.1029/1998GL900033.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E. Array programming with NumPy. *Nature*, 585(7825):357–362, sep 2020. doi: 10.1038/s41586-020-2649-2.

Heimann, S., Kriegerowski, M., Isken, M., Cesca, S., Daout, S., Grigoli, F., Juretzek, C., Megies, T., Nooshiri, N., Steinberg, A., Sudhaus, H., Vasyura-Bathke, H., Willey, T., and Dahm, T. Pyrocko - An open-source seismology toolbox and library. *GFZ Data Services*, v. 0.3, 2017. doi: http://doi.org/10.5880/GFZ.2.1.2017.001.

Heimann, S., Vasyura-Bathke, H., Sudhaus, H., Isken, M. P., Kriegerowski, M., Steinberg, A., and Dahm, T. A Python framework for efficient use of pre-computed Green's functions in seismological and other physical forward and inverse source problems. *Solid Earth*, 10(6):1921–1935, 2019. doi: 10.5194/se-10-1921-2019.

Hofstetter, A., Thio, H. K., and Shamir, G. Source mechanism of the 22 / 11 / 1995 Gulf of Aqaba earthquake and its aftershock sequence. *J. Seismolog.*, 7:99–114, 2003. doi: 10.1023/A:1021206930730.

Hunter, J. D. Matplotlib : A 2D graphics environment. *Comput. Sci. Eng.*, 9(3):90–95, 2007. doi: 10.1109/MCSE.2007.55.

Ide, S. Estimation of radiated energy of finite-source earthquake models. *Bull. Seismol. Soc. Am.*, 92(8):2994–3005, 2002. doi: 10.1785/0120020028.

Ishii, M., Shearer, P. M., Houston, H., and Vidale, J. E. Extent, duration and speed of the 2004 Sumatra-Andaman earthquake imaged by the Hi-Net array. *Nature*, 435(7044):933–936, 2005. doi: 10.1038/nature03675.

Ishii, M., Shearer, P. M., Houston, H., and Vidale, J. E. Teleseismic P wave imaging of the 26 December 2004 Sumatra-Andaman and 28 March 2005 Sumatra earthquake ruptures using the Hi-net array. *J. Geophys. Res. Solid Earth*, 112(B11), 2007. doi: 10.1029/2006JB004700.

Isken, M., Sudhaus, H., Heimann, S., Steinberg, A., Daout, S., and Vasyura-Bathke, H. Kite - Software for Rapid Earthquake Source Optimisation from InSAR Surface Displacement. *GFZ Data Services*, v 0.1, 2017. doi: http://doi.org/10.5880/GFZ.2.1.2017.002.

Jónsson, S., Zebker, H., Segall, P., and Amelung, F. Fault slip distribution of the 1999 Mw 7.1 Hector Mine, California, earthquake, estimated from satellite radar and GPS measurments. *Bull. Seismol. Soc. Am.*, 92(4):1377–1389, 2002. doi: 10.1785/0120000922.

Kanamori, H. and Brodsky, E. E. The physics of earthquakes. *Rep. Prog. Phys.*, 67(8):1429–1496, 2004. doi: 10.1088/0034-4885/67/8/R03.

Kennett, B. and Engdahl, E. Traveltimes for global earthquake location and phase identification. *Geophys, J. Int.*, 105(2): 429–465, 1991. doi: 10.1111/j.1365-246X.1991.tb06724.x.

Khrepy, S. E., Koulakov, I., Al-Arifi, N., and Petrunin, A. G. Seismic structure beneath the Gulf of Aqaba and adjacent areas based on the tomographic inversion of regional earthquake data. *Solid Earth*, 7(3):965–978, 2016. doi: 10.5194/se-7-965-2016.

Kiser, E. and Ishii, M. Back-Projection Imaging of Earthquakes. *Annu. Rev. Earth Planet. Sci.*, 45:271–299, 2017. doi: 10.1146/annurev-earth-063016-015801.

Klinger, Y., Rivera, L., Haessler, H., and Maurin, J.-c. Active Faulting in the Gulf of Aqaba : New Knowledge from the Mw 7 . 3 Earthquake of 22 November 1995. *Bull. Seismol. Soc. Am.*, 89(August): 1025–1036, 1999. doi: 10.1785/BSSA0890041025.

Klinger, Y., Michel, R., and Avouac, J.-P. Co-seismic deformation during the Mw 7.3 Aqaba earthquake (1995) from ERS-SAR interferometry. *Geophys. Res. Lett.*, 27(22):3651–3654, 2000. doi: 10.1029/1999GL008463.

Krüger, F. and Ohrnberger, M. Tracking the rupture of the Mw = 9.3 Sumatra earthquake over 1,150 km at teleseismic distance. *Nature*, 435(7044):937–939, 2005. doi: 10.1038/nature03696.

Lay, T., Ammon, C. J., Kanamori, H., Koper, K. D., Sufri, O., and Hutko, A. R. Teleseismic inversion for rupture process of the 27 February 2010 Chile (Mw 8.8) earthquake. *Geophysical Research Letters*, 37(13), 2010. doi: 10.1029/2010GL043379.

Le Béon, M., Klinger, Y., Mériaux, A. S., Al-Qaryouti, M., Finkel, R. C., Mayyas, O., and Tapponnier, P. Quaternary morphotectonic mapping of the Wadi Araba and implications for the tectonic activity of the southern Dead Sea fault. *Tectonics*, 31(5):1–25, 2012. doi: 10.1029/2012TC003112.

Lefevre, M. *Segmentation des grands décrochements, du cycle seismique à la déformation long terme, exemple de la faille du Levant*. PhD Thesis, Université Paris Diderot, 2018.

Lefevre, M., Klinger, Y., Al-Qaryouti, M., Le Béon, M., and Moumani, K. Slip deficit and temporal clustering along the Dead Sea fault from paleoseismological investigations. *Sci. Rep.*, 8(1):1–9, 2018. doi: 10.1038/s41598-018-22627-9.

Li, X., Jónsson, S., and Cao, Y. Interseismic deformation from Sentinel-1 burst-overlap interferometry: Application to the southern Dead Sea fault. *Geophys. Res. Lett.*, 48: e2021GL093481, 2021. doi: 10.1029/2021GL093481.

Madariaga, R. High-frequency radiation from crack (stress drop) models of earthquake faulting. *Geophys, J. Int.*, 51(3):625–651, 1977. doi: 10.1111/j.1365-246X.1977.tb04211.x.

Meier, M. A., Ampuero, J. P., and Heaton, T. H. The hidden simplicity of subduction megathrust earthquakes. *Science*, 357(6357): 1277–1281, 2017. doi: 10.1126/science.aan5643.

Meng, L., Ampuero, J.-P., Sladen, A., and Rendon, H. High-resolution backprojection at regional distance: Application to the Haiti M 7.0 earthquake and comparisons with finite source studies. *J. Geophys. Res. Solid Earth*, 117(B4):n/a—-n/a, apr 2012. doi: 10.1029/2011JB008702.

Meng, L., Zhang, A., and Yagi, Y. Improving back projection imaging with a novel physics-based aftershock calibration approach: A case study of the 2015 Gorkha earthquake. *Geophys. Res. Lett.*, 43(2):628–636, 2016. doi: 10.1002/2015GL067034.

Minson, S. E., Simons, M., and Beck, J. L. Bayesian inversion for finite fault earthquake source models I-theory and algorithm. *Geophys, J. Int.*, 194(3):1701–1726, 2013. doi: 10.1093/gji/ggt180.

Moral, P. D., Doucet, A., and Jasra, A. Sequential Monte Carlo samplers. *J.R. Statist. Soc. B.*, 68(3):411–436, 2006.

Mustać, M., Hejrani, B., Tkalčić, H., Kim, S., Lee, S. J., and Cho, C. S. Large isotropic component in the source mechanism of the 2013 democratic people's Republic of Korea nuclear test revealed via a hierarchical Bayesian inversion. *Bull. Seismol. Soc. Am.*, 110(1): 166–177, 2020. doi: 10.1785/0120190062.

Okuwaki, R. and Yagi, Y. Role of geometric barriers in irregular-rupture evolution during the 2008 Wenchuan earthquake. *Geophys, J. Int.*, 212(3):1657–1664, 2018. doi: 10.1093/gji/ggx502.

Palo, M., Tilmann, F., Krueger, F., Ehlert, L., and Lange, D. High-frequency seismic radiation from Maule earthquake (M w 8.8, 2010 February 27) inferred from high-resolution backprojection analysis. *Geophys, J. Int.*, 199(2):1058–1077, 2014. doi: 10.1093/gji/ggu311.

Pinar, A. and Türkelli, N. Source inversion of the 1993 and 1995 Gulf of Aqaba earthquakes. *Tectonophysics*, 283(1-4):279–288, 1997. doi: 10.1016/S0040-1951(97)00070-X.

Ribot, M., Klinger, Y., Jónsson, S., Avsar, U., Pons-Branchu, E., Matrau, R., and Mallon, F. L. Active Faults' Geometry in the Gulf of Aqaba, Southern Dead Sea Fault, Illuminated by Multi-beam Bathymetric Data. *Tectonics*, 40(4):1–13, 2021. doi: 10.1029/2020TC006443.

Rössler, D., Krueger, F., Ohrnberger, M., and Ehlert, L. Rapid characterisation of large earthquakes by multiple seismic broadband arrays. *Nat. Hazards Earth Syst. Sci.*, 10(4):923–932, 2010. doi: 10.5194/nhess-10-923-2010.

Schimmel, M. and Paulssen, H. Noise reduction and detection of weak, coherent signals through phase-weighted stacks. *Geophys, J. Int.*, 130(2):497–505, 1997. doi: 10.1111/j.1365-246X.1997.tb05664.x.

Schwarz, G. Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2):461 – 464, 1978. doi: 10.1214/aos/1176344136.

Shamir, G., Baer, G., and Hofstetter, A. Three-dimensional elastic earthquake modelling based on integrated seismological and InSAR data: the M w = 7.2 Nuweiba earthquake, gulf of Elat/Aqaba 1995 November. *Geophys, J. Int.*, 154(3):731–744, sep 2003. doi: 10.1046/j.1365-246X.2003.01978.x.

Shi, X., Wang, Y., Liu-Zeng, J., Weldon, R., Wei, S., Wang, T., and Sieh, K. How complex is the 2016 Mw 7.8 Kaikoura earthquake, South Island, New Zealand? *Sci. Bull.*, 62(5):309–311, 2017. doi: 10.1016/j.scib.2017.01.033.

Spudich, P. and Frazer, L. N. Use of ray theory to calculate high-frequency radiation from earthquake sources having spatially variable rupture velocity and stress drop. *Bulletin of the Seismological Society of America*, 74(6):2061–2082, 12 1984. doi: 10.1785/BSSA0740062061.

Steinberg, A. braunfuss/Palantiri: v0.8 (v0.8), 2021. 10.5281/zenodo.4465171.

Steinberg, A., Sudhaus, H., and Krüger, F. Using teleseismic backprojection and InSAR to obtain segmentation information for large earthquakes: a case study of the 2016 M w 6.6 Muji earthquake. *Geophys, J. Int.*, 232(3):1482–1502, 2022. doi: 10.1093/gji/ggac392.

Steinhaus, H. Sur la division des corp materiels en parties. *Bull. Acad. Polon. Sci*, 1(804):801, 1956.

Sudhaus, H. and Jónsson, S. Improved source modelling through combined use of InSAR and GPS under consideration of correlated data errors: application to the June 2000 Kleifarvatn earthquake, Iceland. *Geophys, J. Int.*, 176(2):389–404, feb 2009. doi: 10.1111/j.1365-246X.2008.03989.x.

Tinti, E., Fukuyama, E., Piatanesi, A., and Cocco, M. A kinematic source-time function compatible with earthquake dynamics. *Bull. Seismol. Soc. Am.*, 95(4):1211–1223, 2005. doi: 10.1785/0120040177.

Vasyura-Bathke, H., Dettmer, J., Steinberg, A., Heimann, S., Isken, M., Zielke, O., Mai, P. M., Sudhaus, H., and Jónsson, S. BEAT - Bayesian Earthquake Analysis Tool. *GFZ Data Services*, v.1.0, 2019.

Vasyura-Bathke, H., Dettmer, J., Steinberg, A., Heimann, S., Isken, M. P., Zielke, O., Mai, P. M., Sudhaus, H., and Jónsson, S. The Bayesian Earthquake Analysis Tool. *Seismol. Res. Lett.*, 91(2A): 1003–1018, jan 2020. doi: 10.1785/0220190075.

Vasyura-Bathke, H., Dettmer, J., Dutta, R., Mai, P. M., and Jónsson, S. Accounting for theory errors with empirical Bayesian noise models in nonlinear centroid moment tensor estimation. *Geophys, J. Int.*, jan 2021. doi: 10.1093/gji/ggab034.

Vasyura-Bathke, H., Steinberg, A., Krüger, F., Guangcai, F., Mai, P. M., and Jónsson, S. InSAR derived surface displacement maps of the 1995 Gulf of Aqaba earthquake [Data set], 2024. 10.5281/zenodo.10462416.

Viltres, R., Jónsson, S., Alothman, A., Liu, S., Leroy, S., Masson, F., Doubre, C., and Reilinger, R. Present-day motion of the Arabian Plate. *Tectonics*, 41:e2021TC00713, 2022. doi: 10.1029/2021TC007013.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, b., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.

Wald, D. J. and Allen, T. I. Topographic slope as a proxy for seismic site conditions and amplification. *Bulletin of the Seismological Society of America*, 97(5):1379–1395, 2007. doi: 10.1785/0120060267.

Wang, D., Kawakatsu, H., Mori, J., Ali, B., Ren, Z., and Shen, X. Back-projection analyses from four regional arrays for rupture over a curved dipping fault: The Mw 7.7 24 September 2013 Pakistan earthquake. *J. Geophys. Res. Solid Earth*, 121(3):1948–1961, 2016. doi: 10.1002/2015JB012168.

Wang, R. A simple orthonormalization method for stable and efficient computation of Green's functions. *Bull. Seismol. Soc. Am.*, 89(3):733–741, 1999. doi: 10.1785/BSSA0890030733.

Wang, R., Lorenzo-Martín, F., and Roth, F. PSGRN / PSCMP — a new code for calculating co- and post-seismic deformation , geoid and gravity changes based on the viscoelastic-gravitational dislocation theory. *Comput. Geosci.*, 32(4):527–541, 2006. doi:

10.1016/j.cageo.2005.08.006.

Wang, R., Heimann, S., Zhang, Y., Wang, H., and Dahm, T. Complete synthetic seismograms based on a spherical self-gravitating Earth model with an atmosphere – ocean – mantle – core structure. *Geophys, J. Int.*, 210:1739–1764, 2017. doi: 10.1093/gji/ggx259.

Wegmüller, U. SAR Processing, Interferometry, Differential Interferometry and Geocoding Software, 1998. presented at EUSAR'98, 25-27 May.

Wesnousky, S. Predicting the endpoints of earthquake rupture. *Nature*, 444(7117):358—-360, 2006. doi: 10.1038/nature05275.

Wessel, P., Smith, W. H. F., Scharroo, R., Luis, J., and Wobbe, F. Generic Mapping Tools: Improved Version Released. *EOS Trans. AGU*, 94(45):409–410, 2013. doi: 10.1002/2013EO450001.

doi:10.26443/seismica.v3i1.1130

SEISMICA

PyOcto: A high-throughput seismic phase associator

Jannes Münchmeyer * [1]

[1]Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, Univ. Gustave Eiffel, ISTerre, Grenoble, France

Abstract Seismic phase association is an essential task for characterising seismicity: given a collection of phase picks, identify all seismic events in the data. In recent years, machine learning pickers have lead to a rapid growth in the number of seismic phase picks. Even though new associators have been suggested, these suffer from long runtimes and sensitivity issues when faced with dense seismic sequences. Here we introduce PyOcto, a novel phase associator tackling these issues. PyOcto uses 4D space-time partitioning and can employ homogeneous and 1D velocity models. We benchmark PyOcto against popular state of the art associators on two synthetic scenarios and a real, dense aftershock sequence. PyOcto consistently achieves detection sensitivities on par or above current algorithms. Furthermore, its runtime is consistently at least 10 times lower, with many scenarios reaching speedup factors above 50. On the challenging 2014 Iquique earthquake sequence, PyOcto achieves excellent detection capability while maintaining a speedup factor of at least 70 against the other models. PyOcto is available as an open source tool for Python on Github and through PyPI.

Production Editor:
Gareth Funning
Handling Editor:
Stephen Hicks
Copy & Layout Editor:
Théa Ragon

Signed reviewer(s):
Sacha Lapins

Received:
October 17, 2023
Accepted:
January 19, 2024
Published:
January 29, 2024

1 Introduction

One of the fundamental tasks in seismology is creating detailed seismicity catalogs. Highly complete catalogs can reveal, for example, spatial migrations, locking patterns, or changes in seismicity rate (González-Vidal et al., 2023; Moutote et al., 2023; Tan et al., 2021). The standard workflow for event detection consists of two steps: phase picking and phase association. The phase picking step identifies the times of seismic phases arrivals in continuous waveforms. The phase association step aims to find consistent sets of picks that can be associated to a seismic source, called an event. This grouping enables downstream analysis steps requiring multi-station data, for example, location or magnitude estimation. In addition, phase association helps to identify and discard spurious picks.

Traditional phase association algorithms often rely on greedy, combinatorial strategies (Johnson et al., 1995). However, these approaches scale poorly with an increasing number of picks. While this has already become a challenge due to the growing number of seismic stations in large-scale deployments, the problem has been supercharged with the advent of highly sensitive, deep-learning-based seismic phase pickers. Deep-learning-based pickers employ neural network models and are trained on millions of manually labeled seismic phase picks. They outperform traditional picking models substantially in terms of sensitivity and pick precision (Zhu and Beroza, 2019; Mousavi et al., 2020; Münchmeyer et al., 2022).

To deal with this flood of phase picks, in recent years, a wave of new phase association algorithms have been published. These approaches range from improved grid-search strategies to complex deep learning architectures. We review the main contributions in the subsequent background section. However, we first discuss the main challenges and performance indicators for seismic phase associators.

The key metric for seismic phase associators is the quality at which they recover seismic events. This includes two aspects: the fraction of events being recovered, i.e. true positive rate or recall, and the fraction of identified events being incorrect, i.e. false positive rate. Usually a tuning parameter can be used to trade-off between those metrics: either a higher recall with a higher rate of false positives or a lower recall with lower false positive rate. The second metric concerns the same questions on pick level: how many picks have been correctly associated and how many picks have incorrectly been associated. Similar trade-offs to the event metrics exist. As ground-truth catalogs for seismicity are not available, seismic phase associators are usually evaluated on synthetic data, i.e., phase picks predicted using travel time calculation and random noise picks. In addition, models are tested qualitatively on real-world example scenarios without ground-truth.

A metric often disregarded is the run time of the algorithms. However, given the ever-growing number of picks, we consider this metric essential to understand the scalability of current algorithms and their applicability to large scale deployments. Run time issues make some of the current associators non-applicable to such deployments, as we show in our examples where some associators did not complete associating a single day of phase picks within 48 hours.

While the recently published associators improve on all of these metrics when faced with large collections of seismic picks, our experiments show that associators

*Corresponding author: munchmej@univ-grenoble-alpes.fr

are still a limiting factor when building seismicity catalogs. This refers to both the precision and recall of events and picks, and the run times, with several associators requiring much more time for association than the phase pickers for picking. For this reasons, we propose PyOcto, a novel **Py**thon-based associator inspired by the **Octo**tree data structure. PyOcto is based on the idea of dividing space-time into potential origins. It achieves fast run times by only exploring promising origin regions, making it a high-throughout phase associator. PyOcto is available as an open source code with a range of different input and output interfaces for easy use.

2 Related work

Before describing the PyOcto architecture, we introduce the most popular novel seismic phase association methods published within the last years. All described algorithms rely on first arriving P and S phase picks without taking into account later phases. REAL (Zhang et al., 2019) is an optimized grid-search algorithm. Instead of searching a full space-time grid, REAL is based on the assumption that a station close to the event will record the first P pick. Starting with one P pick, a grid search is performed in a volume around the picking stations. This reduces the search space from the whole study area to a smaller volume. In addition, it removes the time dimension from the search, as the approximate origin time for each potential origin can be inferred from the starting pick. REAL can use homogeneous and 1D velocity models.

HEX (Woollam et al., 2020) is a hyperbolic phase associator. Assuming a homogeneous velocity model, it postulates that the picks of one event need to occur on a hyperbola. HEX uses the probabilistic RANSAC algorithm to fit such hyperbolas to the picks. In this algorithm, random candidate sets of picks are drawn and a hyperbola is fit. If the hyperbola contains sufficiently many picks, an event is declared.

GaMMA (Zhu et al., 2022) is based on a similar assumption of a hyperbolic moveout but uses a different optimisation scheme. The method interprets the picks as a Gaussian mixture with each event a different mixture component. GaMMA uses an expectation-maximization (EM) algorithm for optimizing the clusters. As run times for the EM algorithm grow substantially superlinearly with the number of picks, GaMMA uses DBSCAN (Ester et al., 1996) to group picks before applying the EM algorithm to each cluster. GaMMA was originally published with a homogeneous velocity model but has later been extended to support 1D models too. Ross et al. (2023) proposed Neuma, a generalisation of GaMMA using an Eikonet (Smith et al., 2020) to enable arbitrary 3D velocity models instead of the homogeneous velocity model.

In addition to these optimization based algorithms, several deep learning models have been proposed for phase association. PhaseLink (Ross et al., 2019) uses a recurrent neural network applied to pick times, phase type and station locations to identify pairwise associations between picks. It then employs an aggregation step to infer consensus sets of matching phases that correspond to event detections. GENIE (McBrearty and Beroza, 2023) uses a Graph Neural Network. Similar to PhaseLink, GENIE uses the arrival time, phase type and station location as inputs. In contrast to PhaseLink, GENIE treats all picks jointly and outputs the full association result from the neural network. Both GENIE and PhaseLink are trained on synthetic data generated using 1D velocity models. The training step needs to be conducted once for each target region, afterwards the models can be applied to arbitrary amounts of data. Deep learning methods differ fundamentally from classical approaches as their actual application relies on detecting patterns rather than applying some travel time based approach. While these patterns are clearly based on travel times, this can make the models harder to debug and interpret. On the other hand, it brings advantages such as the ability to identify which collections of picking/non-picking stations might be reasonable.

3 Methods

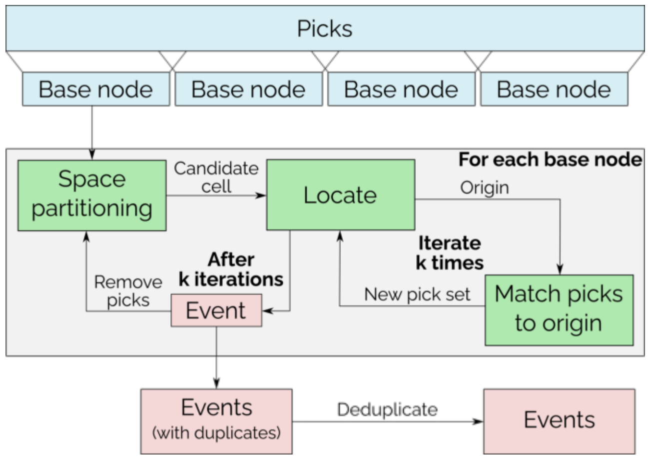


Figure 1 Schematic view of the full PyOcto pipeline. The picks are split by time into base nodes. For each base node, the grey box is executed. Several of these boxes can be executed in parallel. Within each box, the space partitioning algorithm (see Figure 2) and the localisation/pick matching steps are conducted. Events are output and finally deduplicated.

In the following sections we present the PyOcto associator. We start with the core algorithms and then discuss details, optimisations and implementation details. A schematic overview of the full associator is provided in Figure 1. Throughout the description we add the parameter names used in the implementation in italics in brackets to allow easier cross-referencing.

3.1 Core algorithm

PyOcto is based on partitioning space-time into cells. The key idea is to mimic a grid-search associator while only looking at "useful" grid cells. We achieve this by using a data structure inspired by an octotree with an additional time axis. The data structure consists of a collection of 4D volumes (3D in space, 1D in time), that we will

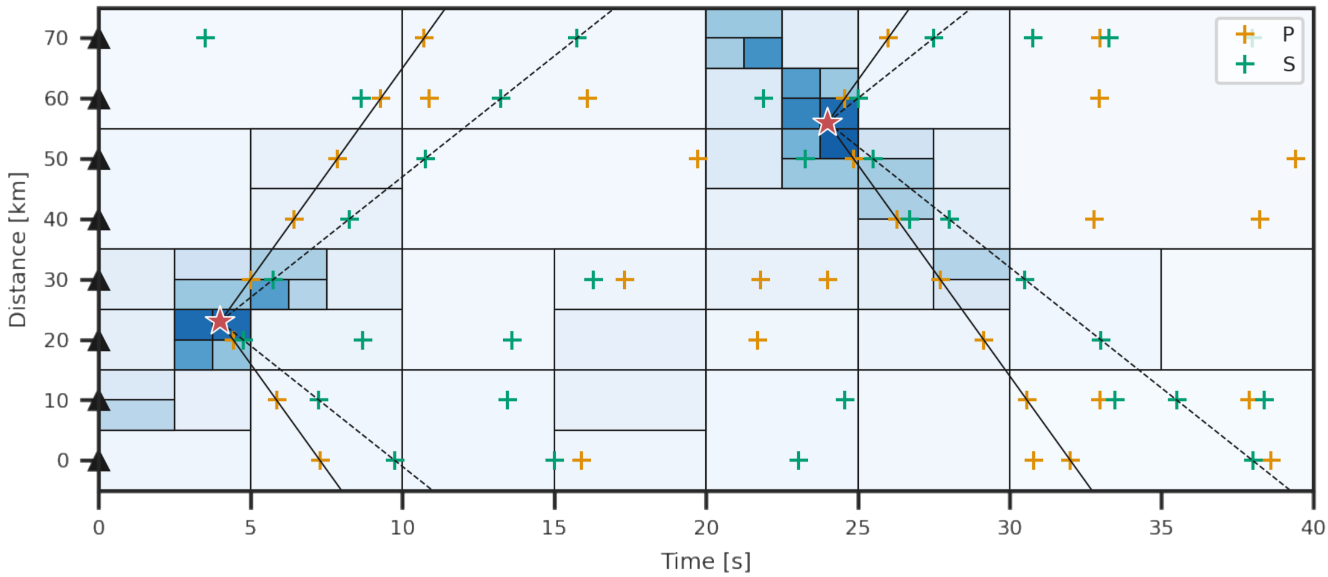


Figure 2 Schematic view of the gridding scheme with only one spatial dimension and the time dimension. Picks are indicated by crosses, the station locations are marked on the left by black triangles. Two events are contained, marked by red stars with P (solid) and S (dashed) moveout shown in black. The background shows the gridding with each cells shading corresponding to the number of picks per area . Only cells with at least 6 matching P picks and 6 matching S picks were explored. For each area, only the smallest cell explored is shown, i.e., all larger cells explored before in the same region are not visualised.

call nodes in the following to highlight the resemblance of a tree data structure. We show a simplified version of this with only one space axis in Figure 2. Each volume/node V is associated to the list of picks $picks(V)$ that could have originated from the node. More formally, let V be a node and (s, t) a pick at station s at time t.[1] We write

$$(s, t) \in picks(V) \Leftrightarrow \exists (x_0, t_0) \in V : t_0 + tt(x_0, s) = t + \epsilon \quad (1)$$

with $tt(x_0, s)$ the travel time from the origin x_0 to the station s. We include an ϵ to indicate that the equation only needs to hold up to a given uncertainty (*tolerance*). This uncertainty takes into account inaccuracies in the velocity model and the pick times. For simplicity, PyOcto uses a fixed tolerance threshold that is identical for all picks.

There are two crucial insights about the definition of picks belonging to a node. First, while for each pick there exists a location/time in the node where it could have originated, this location/time might be different for each pick. Therefore, a set of picks originating from a node is not a sufficient condition for associating these picks into an event. This becomes obvious when looking at very large nodes. On the other hand, it is a necessary condition, i.e., if there is an event with sufficiently many picks in the dataset, there must be a node that contains all these picks. Second, the assignment of picks to nodes is not unique. A pick might be contained in multiple nodes, even if these nodes are disjoint. However, only few nodes will contain enough picks to produce an event. The key idea of PyOcto is to cleverly identify these nodes.

[1] For simplicity we omit the phase of the pick here. The inclusion of phase type is natural and only involves taking different travel time models for P and S waves.

PyOcto starts with a large node spanning the whole study area and a long time. All picks recorded during this time (with adjustments for boundary effects) can be assigned to the node. We initialize a list of active nodes with this node. The association then repeatedly takes the active node with the highest number of picks and performs one of the following actions:

- if the node can not create an event anymore: discard node
- if the node is small enough: try creating an event
- otherwise: split the node and add children to the list of active nodes

We use a priority queue for the list of active nodes to efficiently retrieve the node with the highest number of picks. In the following, we describe the different actions.

Splitting a node: The most common action is splitting a node. For this action, we split the node V into two disjoint children V_1 and V_2, such that $V = V_1 \cup V_2$. We split V in half along the coordinate axis in which V has the largest extent. To compare the time axis, we multiply it with a constant velocity, by default 5 km/s.

We then build the sets $picks(V_1)$ and $picks(V_2)$ by iterating over all candidates in $picks(V)$. This check can easily be performed using equation (1). As noted before a pick can be assigned to both of these sets at the same time.

Discarding a node: Essential for the high performance of PyOcto is to discard nodes early if they can not produce an event anymore. For this, we use the following criteria:

- minimum number of total picks (*n_picks*)
- minimum number of P picks (*n_p_picks*)
- minimum number of S picks (*n_s_picks*)
- minimum number of stations with both P and S picks (*n_p_and_s_picks*)

All thresholds are configurable and should be adjusted to the dataset. As a subvolume can never contain more picks than the parent node, once a node violates any of these criteria it can not create an event anymore and can be discarded.

Creating an event: If a cell is smaller than a predefined threshold along all axes (*min_node_size*), PyOcto tries to create an event. For this, we locate an event based on all picks in a cell. The full localisation procedure is described in Section 3.2. We then identify whether all picks fit the determined location and remove potential outliers. These outliers might occur as not all picks in the node need to necessarily stem from the same source location/time. In addition, we scan all other picks to identify if further picks are consistent with the list of picks. This operation can be performed efficiently using a binary search in time. We add these picks to the list of picks. This procedure is repeated multiple times (*refinement_iterations*), by default 3, to stablise the event. If at any point in this iteration the picks do not fulfill the conditions for nodes outlined above, the event creation is stopped as unsuccessful.

Even though the node already gives a preliminary location and station set, the location procedure is required for multiple reasons. First, while the node groups a candidate set of picks, there is no guarantee that all of these can be associated to a common origin. Second, the optimal location for a set of picks does not necessarily need to fall within the node, in particular, because the same set of picks can be contained in multiple nodes. This is also the reason why it might be possible to associate additional picks to the location. While traversing the nodes by number of picks makes it likely to select nodes already containing the majority of picks for an event, this can not be guaranteed in face of spurious picks.

In contrast to some other associators (e.g., GaMMA), PyOcto can not use amplitude information for association. However, obtaining accurate amplitudes for events at low signal-to-noise levels, as for the majority of events detected with deep learning, is challenging. From our anecdotal experiments on real data, we did not see a major advantage from the use of amplitude information.

3.2 Localisation procedure

To identify the most likely origin for a set of picks, we use the equal differential-time (EDT) loss (Lomax et al., 2000). Compared to an L2 loss on the travel time residual, the EDT loss has two advantages. First, it is independent of the origin time, thereby reducing the search space. Second, it is more stable against outlier picks. As we expect outliers to be contained in our pick set, this is a useful property for our application.

To find the minimum of the EDT loss, we use a greedy algorithm. A greedy algorithm is a heuristic making locally optimal decisions. While such algorithms do not necessarily find the global optimum, they usually show excellent runtimes. Starting with the whole study volume, we split the volume in half k times (*location_split_depth*) into 2^k subvolumes. For each subvolume, we calculate the EDT loss at the volume center. From the volume with the lowest EDT loss, we go up l splits (*location_split_return*). This volume, with a size of $1/2^{k-l}$ of the original volume is used as the new start for the search and we repeat the splitting and search procedure. We iterate this step until the volume reaches a predefined size (*min_node_size_location*).

This greedy algorithm has a trade-off between accuracy and runtime. When splitting the volume into only few pieces and only using a low l, this leads to low runtime but potentially suboptimal minima. On the other hand, too fine splitting in each step will increase runtimes at virtually no gains in location accuracy. We set the default to $k = 6$ and $l = 4$, but make the parameter individually configurable. We note that insufficient values for k and l can lead to striping artifacts, i.e., locations at the edges of larger volumes caused by insufficient sampling.

3.3 Velocity models

At its core, PyOcto relies on travel times. These travel times need to be obtained from seismic velocity models. Two types of queries occur in the PyOcto algorithm. First and most commonly, volume queries of type $(s, t) \in picks(V)$, i.e., identifying if a pick can originate from a volume. Second, for the localisation algorithm, traditional travel times between the proposed origin and the station are required. Both of these queries will be executed in very high numbers and therefore need to be implemented efficiently.

PyOcto implements two velocity models, a homogeneous model and a 1D layered model. For the homogeneous model, we assume constant P and S velocities. To solve the volume query, we identify the earliest and latest times a pick from the volume could arrive at the station. The earliest time is achieved by the earliest origin time in the volume plus the travel time to the closest point in the volume. Similarly the latest time can be derived using the point with the highest distance to the station. For a homogeneous velocity model, the derivation of the travel times from a fixed origin are trivial using Pythagoras theorem. Both queries run in constant time.

For the layered velocity model, we use a precalculation step to substantially improve performance. First, we calculate P and S arrival times on a dense grid using an eikonal solver. This step takes a few seconds but only needs to be run once. For extracting travel times we run 2D bilinear interpolation between the 4 closest grid nodes. For the area queries, i.e., if a pick can result from a volume we use the observation that for a 1D velocity model, the shortest travel time must be at the closest epicentral distance and the longest travel time at the furthest. However, it is not a priori clear at which depth these times occur. Potential candidates are the shallowest and deepest points of the queried depth interval, plus all local extrema within the depth interval. Local extrema are regarded in depth direction, e.g., at a fixed distance a depth is a local minimum if the travel times both directly above and below it are larger. To efficiently query the local extrema, we cache all local extrema at each distance. As for typical velocity models each distance has at most a handful of local extrema, they can simply all be checked when necessary. In addition, to correct for station elevation, we add an elevation correction based on a constant velocity and vertical incidence. While this is an approximation, errors are negligble for association purposes.

PyOcto does not support 3D velocity models as performing efficient, i.e., constant run time, volume queries as required for the splitting algorithm is nontrivial. This is identical to most common algorithms, that are limited to homogeneous or 1D models. In contrast, deep learning models are able to use arbitrarily complex models (Ross et al., 2019; McBrearty and Beroza, 2023). PyOcto supports different velocity models between the splitting and the localisation step. In principle, it would be easy to extend at least the localisation step to 3D models. However, we have not tested this and only expect substantial improvements for regions with velocity structures strongly deviating from a layered model.

PyOcto supports station terms, i.e., constant time offsets for phase arrivals at a station, which can occur due to local structure. We implement additive station terms, i.e., the station term is added to the predicted travel time from the velocity model. This is the same sign convention as used by NonLinLoc (Lomax et al., 2000). Station terms are not determined dynamically but have to be defined before running the association. However, they can be obtained by iteratively running PyOcto and inferring station terms from the residuals of the previous run.

For efficient calculation of distances, PyOcto relies on local coordinate transforms. By default, we suggest transverse Mercator projections. The transformation from latitude and longitude values to local coordinates needs to be performed only once before the association step. While distance measures will become inaccurate for very large study areas, we did not observe any issues in our case studies with diameters up to ~ 1500 km.

3.4 Initialisation

As described in the introduction of the algorithm, the association starts with a node spanning the whole study area. In principle, this node could also span the whole study time. However, in practice this is suboptimal because it will require several costly splits along the time axis that are mostly trivial. Instead, we do not start with a single node but with a list of base nodes.

Each base node spans the whole study area but only a part of the time. For this, we split the time into regularly spaced segments, by default 20 minutes long (*time_slicing*). The segments overlap by a short buffer time window (*time_before*). Each segment is then filled with all picks that originate during the segment. With a buffer time roughly equivalent to the maximum travel time through the study area, this ensures that each event is contained completely in at least one base node.

As two subsequent base nodes might both contain most picks for one event, the early splitting might lead to duplicate events. For this reason, we deduplicate the events after all base nodes have been processed. To avoid issues from inaccurate estimates of location or origin time, we base the deduplication exclusively on the set of picks. If two events share more than a fixed number of picks (*max_pick_overlap*), the event with fewer picks is discarded. We note that we allow some level of intersection to avoid discarding nonidentical events. This might lead to some picks being assigned to multiple events, which is not possible for events within one time slice.

3.5 Optimisations

While the splitting algorithm with early stopping is a solid basis for an efficient algorithm, several points need to be taken into account that might affect runtime. Before going into details, we review the general runtime principles. While a formal analysis of algorithm complexity is difficult, we can make several observations. First, run time crucially depends on the number of nodes processed. It is therefore essential to stop the processing of each branch of the search tree as early as possible. Second, location procedures are expensive as they require many travel time queries. They should therefore not be triggered too often. Based on these observations, we define multiple optimisations.

As a first optimisation, PyOcto keeps track of all picks that have already been assigned to events. Once a pick has been assigned to an event, it is not considered anymore and removed from all nodes. Without these picks, the adjacent nodes most likely will not fulfill the necessary minimum number of picks. This step substantially improves runtime, as events will usually produce many adjacent nodes with high numbers of picks which do not need to be processed multiple times.

The second observation treats the case of a group of picks that can not be associated to a common origin. This happens if the origin determined from the group of picks does not correspond to sufficiently many picks, i.e., depends on the tolerance for matching picks and the required number of picks for an event. As trying to create an event from these picks does not yield a consistent origin, these picks are not marked as used. However, often many neighboring cells contain the same set of picks, leading to repeated but useless tries of locating the same set of picks. To mitigate this situation, we cache all sets of picks that have been processed as candidate sets for localisation. If a set has been processed before, it will be skipped in the next try. Note that this optimisation only works because the location search depends only on the pick set but not on the location of a node.

The last optimisation is relevant in the case of a large number of stations with spurious picks. With a growing number of stations, it becomes likely that a set of distant stations by chance produces picks that can be associated. This does not only lead to false detections but also substantially increases run time. At the same time, these false events are easy to identify manually because of the inconsistent pick pattern, i.e., the existence of many non-picking stations between the picking stations.

To remove this issue, we introduce two distance conditions, a relative and an absolute condition. The absolute condition is a simple cutoff on the maximum distance between stations and sources for the space partitioning (*association_cutoff_distance*). This condition excludes picks too far from a given cell when checking if the pick could have originated there. However, in the localisation and pick matching step, all picks are taken into account, ensuring that the output contains all associated picks even at larger distance. This condition is most helpful in large, homogeneous networks and in networks without large amounts of out-of-network events. By default, no maximum distance is set to account for the different scales at which PyOcto might be applied.

For the case of inhomogeneous networks or networks with substantial out-of-network events, we introduce a relative distance condition, based on the assumption that it is unlikely for a station to detect an event if many closer stations did not detect it. For every distance from a volume, we can calculate the fraction of stations within this distance that have at least one pick compared to the total number of stations. We then identify the maximum distance where this fraction is still above a predefined threshold (*min_pick_fraction*). All picks at stations above this threshold are removed. As the nodes have a spatial extent, for each station we choose the distance maximizing the number of retained picks. This means that for stations with picks we use the minimum distance to the node while for all other stations we use the maximum distance. The default value for this threshold is 0.25, i.e., allowing for many close stations without picks.

While this optimization yields substantial runtime improvements for datasets with high numbers of stations, it comes at a cost. To check the condition, at every node the distance to all existing stations needs to be calculated. For small deployments, associations by chance are anyhow unlikely, rendering the additional runtime mostly useless. The optimisation can therefore be deactivated.

Lastly, PyOcto uses a memory protection strategy. As PyOcto processes nodes ordered by their number of picks, it needs to always hold a queue of active nodes. This can, in the worst case, degrade into a breadth-first search, which is very memory intensive. Therefore, once the total number of nodes exceeds a predefined threshold (*queue_memory_protection_dfs_size*), PyOcto processes the next nodes using depth-first search. This is highly memory efficient, as only the current call stack needs to be kept in memory. At the same time, this can lead to increased run times. We speed the search up by still always traversing greedily into the larger of the two children of a node. In our experiments, the memory protection was only required for very large sets of picks in short times (≫100,000 picks per day).

3.6 Implementation

PyOcto is implemented in Python and C++. The interface of PyOcto is implemented in Python to provide an accessible interface in a common scripting language. Inputs and outputs are passed as Pandas data frames. PyOcto has a slim set of dependencies. The backend of PyOcto is implemented in C++. The functions are natively embedded into Python using pybind11. The association function is parallelised using pthreads. Parallelisation is achieved by assigning base nodes to threads. This causes very low synchronisation overhead as only the base node queue and the event list are shared between threads. The list of used picks is not shared between threads, instead events are deduplicated at the end of the association step. By default, PyOcto uses all available threads. However, the thread count can be set manually (*n_threads*).

To allow an easy experimentation with PyOcto, the software implements several compatibility interfaces:

- a function to read the input format from GaMMA (Zhu et al., 2022)
- a function to read the input format from REAL (Zhang et al., 2019)
- a function to process SeisBench picks (Woollam et al., 2022)
- a function to use obspy Inventory objects as input (Beyreuther et al., 2010)
- an output interface for NonLinLoc (Lomax et al., 2000)
- an automated selection strategy for local coordinate transforms

PyOcto is available as open source code under MIT license, a permissive open-source license. Pre-built wheels for Linux, Mac OS, and Windows are available on PyPI and can be installed using pip (see Data and Code availability for details). The PyOcto code is modular to allow for easy extension. Such extensions could, for example, include pick-specific uncertainties or more accurate calculation of topography corrections.

4 Benchmark on synthetic catalogs

4.1 Setup

To quantitatively assess the quality of PyOcto, we test it on synthetic catalogs. We use two complementary scenarios: (i) uniformly distributed seismicity in a shallow layer; (ii) realistic subduction zone seismicity (Figure S1). We compare the proposed PyOcto algorithm to two established associators: GaMMA and REAL. We choose these algorithms as they have well-documented, open-source implementations and have both been used in numerous application cases already (Wilding et al., 2023; González-Vidal et al., 2023; Tan et al., 2021; Liu et al., 2020). We do not compare PyOcto against any deep learning associator, as optimizing these associators requires substantially more parameter choices and

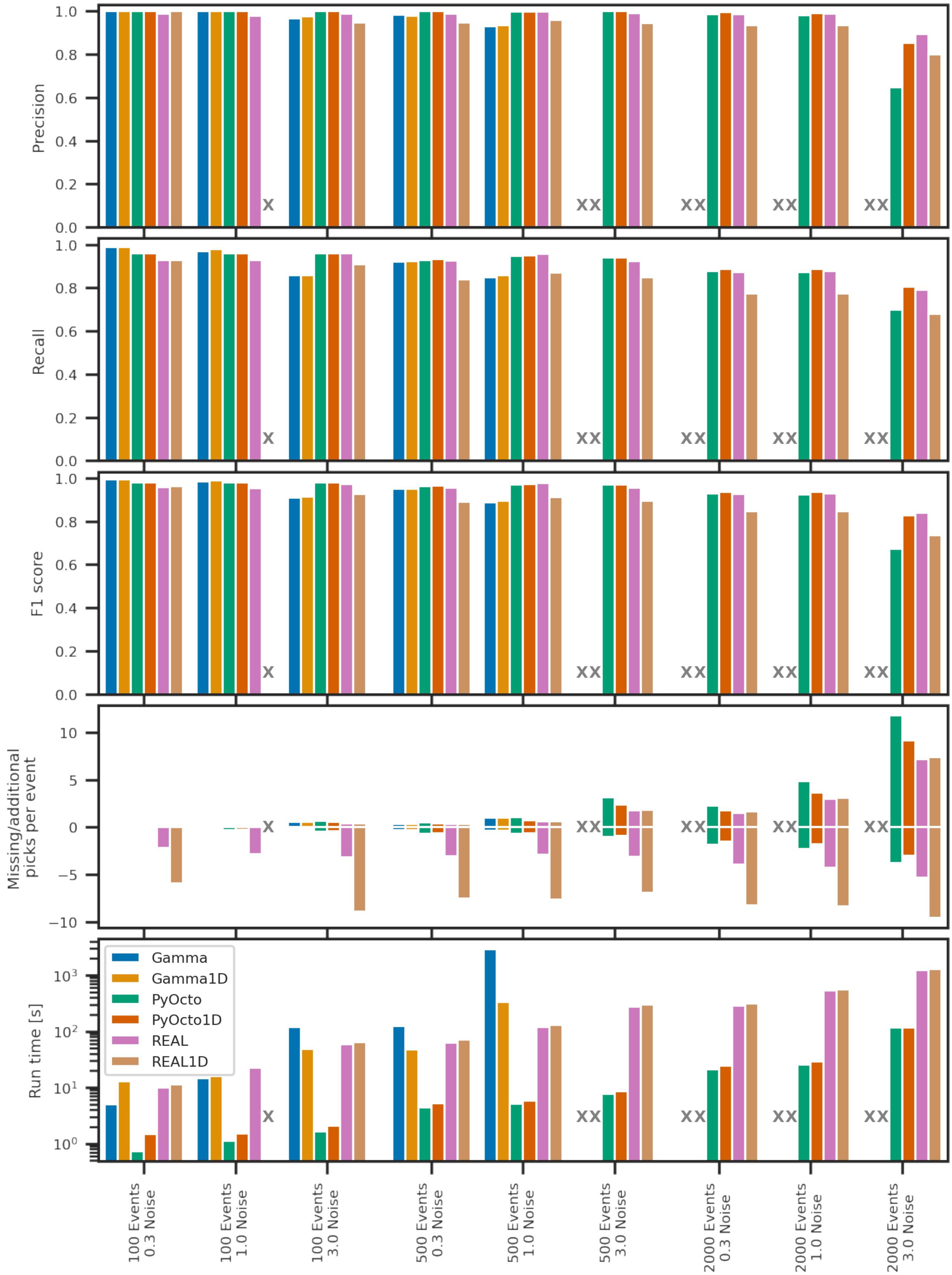


Figure 3 Synthetic evaluation of the different associators in the shallow seismicity scenario. Each associator is indicated by a color. For the missing/additional picks, missing picks are indicated with a bar below 0, additional picks with a bar above 0. Missing results due to exceeded runtimes are indicated by grey Xs. A result for REAL 1D with 100 events and 1.0 noise is not available as the model reproducibly crashed with a segmentation fault. All results in numerical form are reported in Table S4.

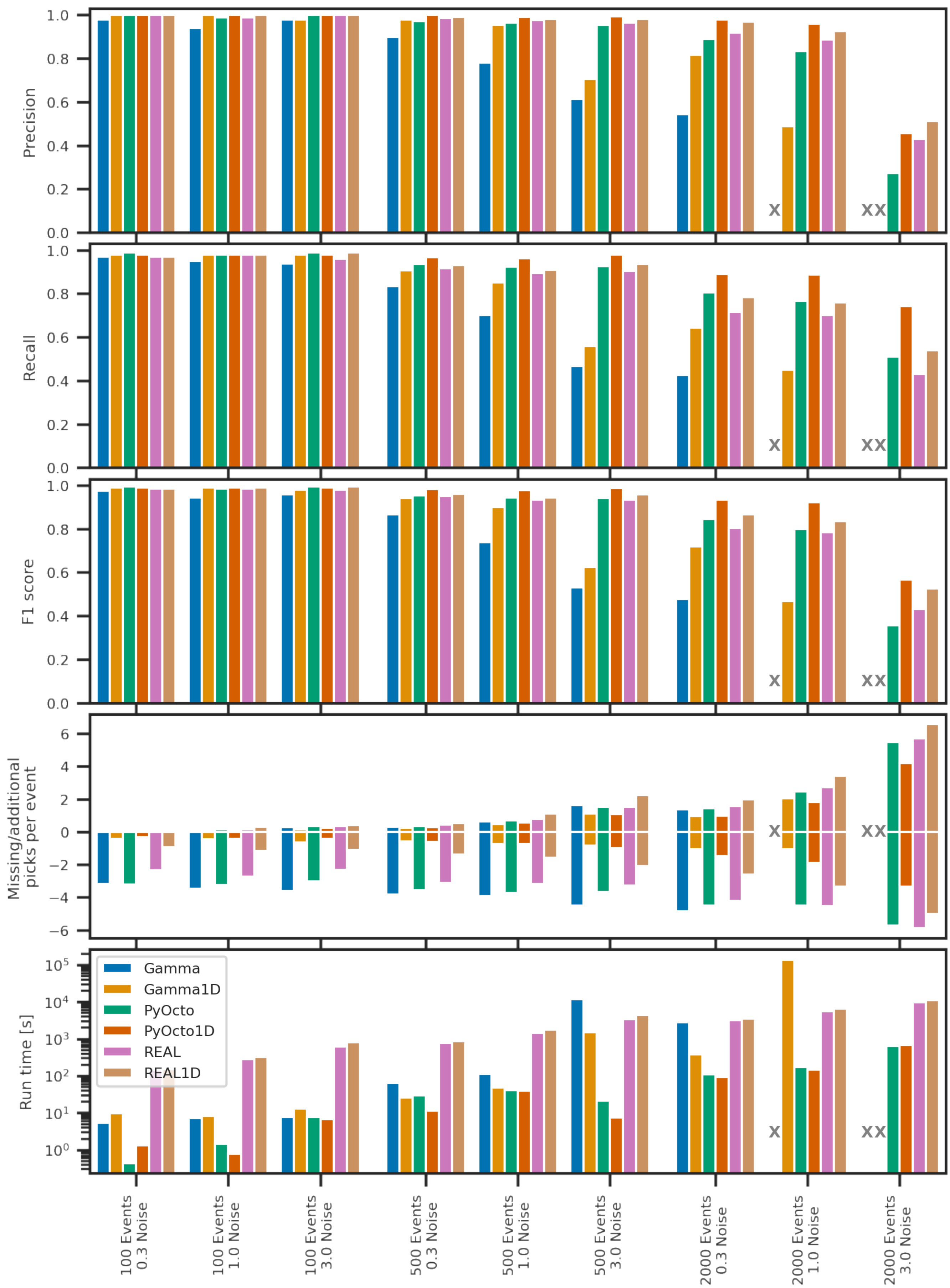


Figure 4 Synthetic evaluation of the different associators in the subduction scenario. For further details see the caption of Figure 3. All results in numerical form are reported in Table S4.

Table 1 Dataset statistics for the subduction scenario. We do not differentiate between P and S picks as both are generated in almost equal number. The picks per station include the noise picks.

Events	Noise	Event picks	Noise picks	Total picks	Avg. picks per event	Avg. picks per station
100	0.3	2,241	672	2,913	22.41	145.65
100	1.0	2,331	2,331	4,662	23.31	233.10
100	3.0	2,142	6,426	8,568	21.42	428.40
500	0.3	11,414	3,424	14,838	22.83	741.90
500	1.0	11,194	11,194	22,388	22.39	1119.40
500	3.0	10,818	32,454	43,272	21.64	2163.60
2,000	0.3	45,544	13,663	59,207	22.77	2960.35
2,000	1.0	45,011	45,011	90,022	22.51	4501.10
2,000	3.0	45,213	135,639	180,852	22.61	9042.60

a fair comparison is therefore harder to guarantee. Note that this study is not intended as a full-scale benchmark of seismic phase associators as this would be out of scope for the paper. Instead, we restrict ourselves to this smaller-scale case study.

Both scenarios use the same procedure for data generation. Each test case consists of one day of seismicity with a predefined number of events and a predefined noise rate. For each event, we draw a source time uniformly within the day and draw a location and a magnitude from the distributions described below. Based on the magnitude and hypocentral distance, we estimate detection probabilities at each station. From these probabilities we randomly select whether a station has a P and an S arrival using correlated Bernoulli variables with correlation 0.5 between the two phases (see supplement Section S1). We predict travel-times using a 1D velocity model from Graeber and Asch (1999). To each individual travel-time we add a Gaussian random normal variable with a standard deviation of 0.4 s, but at least 1 % of the total travel time. Finally, we add noise picks not associated to any event to the data set. The number of noise picks is defined as the product of the number of event picks times the user-defined noise rate. For each pick, the phase, time and station are drawn according to a uniform random distribution. We use event numbers of 100, 500 and 2000, and noise rates of 0.3, 1.0 and 3.0.

We compare PyOcto to GaMMA and REAL. For each model, we manually selected reasonable parameters. All parameters are reported in Tables S1, S2, and S3. For each associator, we report results for the versions with homogeneous velocity models and 1D layered velocity models. We provide the associators with the same velocity model we used for data generation. We therefore expect slightly too optimistic performance results for the 1D models, however, the comparison between these models should still provide reasonable results.

For all associators we require at least 10 picks for an event detection. We furthermore require at least 4 stations with both P and S pick for REAL and PyOcto. We do not enforce the last condition for GaMMA as the option is not implemented. We ensure that all events in our synthetic catalogs fulfill these conditions.

We evaluate the associators based on 6 metrics: precision, recall, F1 score, missing picks per event, incorrectly associated picks per event, and run time. Precision is the fraction of cataloged events among all detections. Recall is the fraction of events detected among all cataloged events. F1 score is the harmonic mean of precision and recall. To calculate these metrics, we define matches between cataloged and detected events through their picks. A cataloged event A and a detected event B are considered a match if at least 60 % of the picks of A are also picks of B and vice versa. We use a pick-based matching instead of a location- and time-based matching as it is more stable for high event densities.

We execute the test on 16 virtual CPU cores with 8 physical cores and 64 GB main memory. We measure runtimes from the invocation to the output of the models. We do not measure data-independent preprocessing steps such as velocity model building as these steps only need to be executed once in an application scenario. As an exception, the times for GaMMA with 1D velocity model includes the time for the eikonal solver ($\sim$3 s), as the step can not be executed separately. Exact machine configurations can vary slightly between tests, therefore the reported runtimes should be interpreted rather as an indication than an exact measure. We limit the total aggregated runtime of all tests per associator to 48 h. All tests not finished at this point are reported as missing.

4.2 Uniform shallow seismicity

As a first scenario, we study shallow seismicity. We use 100 stations arranged in a 10x10 grid with a station spacing of $0.2° \times 0.2°$. Event locations are randomly distributed within the network with a depth up to 30 km. No out-of-network events are generated. Magnitudes are generated from a Gutenberg-Richter distribution with a minimum magnitude of 0.5 and $b = 1$. Dataset statistics are reported in Table 2.

Figure 3 shows the performance metrics for the shallow scenario. Full results in numerical form can be found in Table S4. PyOcto and REAL obtained results for all cases with both the homogeneous and the 1D velocity model. GaMMA did not provide solutions for the cases after 500 events and a noise factor of 1.0 as the computation did not finish within the 48 h time limit.

In all cases, PyOcto achieves the highest F1 score

Table 2 Dataset statistics for the shallow seismicity scenario. We do not differentiate between P and S picks as both are generated in almost equal number. The picks per station include the noise picks.

Events	Noise	Event picks	Noise picks	Total picks	Avg. picks per event	Avg. picks per station
100	0.3	4,047	1,214	5,261	40.47	52.61
100	1.0	4,894	4,894	9,788	48.94	97.88
100	3.0	5,257	15,771	21,028	52.57	210.28
500	0.3	25,658	7,697	33,355	51.32	333.55
500	1.0	24,525	24,525	49,050	49.05	490.50
500	3.0	23,646	70,938	94,584	47.29	945.84
2,000	0.3	101,614	30,484	132,098	50.81	1320.98
2,000	1.0	98,680	98,680	197,360	49.34	1973.60
2,000	3.0	94,710	284,130	378,840	47.35	3788.40

or a result within 0.01 F1 score of the best model. The 1D model slightly outperforms the homogeneous model. REAL with a homogeneous model achieved a slightly worse performance, followed by REAL with a 1D model. GaMMA shows a clear degradation in F1 score with growing number of event or noise picks but still achieves good performance (F1 $\geq$ 0.89) for all cases where solutions were obtained. For the case with 2000 events and a noise factor of 3.0, REAL (homogeneous, 0.84) performs best, closely followed by PyOcto (1D, 0.83), REAL (1D, 0.74), and PyOcto (homogeneous, 0.67). We suspect that REAL shows slightly better performance here because the actual grid search is less affected by noise picks than the approximation using space partitioning used in PyOcto. We note that this case is extremely challenging with each station reporting on average one pick every 23 s.

Up to 500 events and a noise rate of 1.0, PyOcto (1D and homogeneous) and GaMMA (1D and homogeneous) are very exact in terms of picks with few additional or missed picks. In contrast, REAL (homogeneous) misses roughly 3 picks per event, REAL (1D) between 5 and 10. While we are not fully certain about the missed picks, we assume it is because REAL discards picks based on the ratio between station residuals and event residuals, i.e., a low average pick residual for an event will lead to discarding picks with higher residuals even if their absolute value is not excessively high. We note that the number of missed picks for REAL could likely be reduced through targeted parameter tuning. For configurations with high numbers of events, in particular, in conjunction with high noise, REAL and PyOcto both include false picks with the events. PyOcto includes more false picks than REAL, again likely related to selection criteria. The homogeneous version of PyOcto produces about 1.5 times as many false picks as the 1D variant, likely because of the overall higher tolerance value necessary to mitigate the less accurate velocity model.

In terms of run time, PyOcto substantially outperforms GaMMA and REAL in all cases. The run time factor between PyOcto and the next-fastest methods exceeds 10 in almost all cases, often even reaching factors of 20 and above. Run times for the homogeneous and the 1D velocity model for PyOcto are almost identical in all cases. We suspect that while the travel time lookup for the 1D velocity model is slightly more expensive than for the homogeneous model, this effect is offset by more focused origins from the better travel times, leading to fewer nodes that need to be explored. Probably owing to the same better focus, for GaMMA the 1D model is usually faster than the homogeneous model by a factor of 3 to 10.

4.3 Subduction zone

For the subduction zone scenario, we base our catalog on the IPOC network (GFZ German Research Centre For Geosciences and Institut Des Sciences De L'Univers-Centre National De La Recherche CNRS-INSU, 2006) and the catalog by Sippl et al. (2018). We chose the deployment and the catalog as a typical example of a well-instrumented, highly active subduction zone with diverse seismicity. We draw event locations and event magnitudes independently from the catalog. We use the IPOC stations, in total 20 stations. The study area covers approximately 5° North-South and 3° East-West up to a depth of 200 km. Out-of-network seismicity is located up to 1° from the network. This is a typical challenge for associators in subduction zones where offshore events will occur substantially outside the network. Dataset statistic are reported in Table 1.

The results in the subduction scenario largely mirror the ones from the shallow scenario but with nuanced differences that we point out in the following (Figure 4, Table S4). First, the difference between 1D and homogeneous models is more pronounced with 1D models clearly outperforming homogeneous models in terms of F1 score. Furthermore, the homogeneous models (GaMMA, REAL, and PyOcto) consistently miss around 2.5 picks per event. This highlights that the assumption of a homogeneous velocity model is insufficient for subduction zones. Nonetheless, PyOcto and REAL with homogeneous velocity model still achieve F1 scores consistently above 0.93 for cases with up to 500 events. In contrast, GaMMA performs clearly worse than the other models already at 100 events per day and it substantially degrades further above. This happens for GaMMA with a homogeneous velocity model and with a 1D model. We suspect that the degradation for GaMMA is related to the optimisation strategy that is susceptible to local minima. At high numbers of picks, the loss landscape will look very rough, leading to unfavourable convergence properties. This is more pronounced for the subduction scenario than the shallow scenario, as local minima are particularly likely among the depth axis.

Second, among the 1D models, PyOcto outperforms REAL and GaMMA more clearly than in the shallow case. It consistently exhibits a higher F1 score and lower numbers of missed and false picks. Even at 2000 events with a noise rate of 3.0 (with a pick per station on average every 9.5 s), PyOcto still achieves an F1 score of 0.57.

Third, run time differences are even more pronounced with PyOcto outperforming REAL often by a factor of 1000. This is caused by the larger search grid required by REAL to handle the depth and the out-of-network events. We note that we already reduce the impact of the larger grid size for REAL by using a larger grid spacing for the subduction scenario. In contrast, PyOcto can easily handle large search domains due to its splitting approach that scales logarithmically with volume. For the subduction scenario, PyOcto with a 1D model in most cases only needs about half the time of PyOcto with a homogeneous velocity model. This suggests that the more accurate velocity model leads to fewer nodes needing to be explored. GaMMA with a homogeneous velocity structure shows competitive run times compared to PyOcto and REAL for cases with 100 events but run times substantially exceed the ones of REAL (and thus even more PyOcto) at 500 events and above. With a 1D velocity model, GaMMA runtimes are close to the ones of PyOcto as well for 500 events with noise levels of 0.3 or 1.0, however, at the cost of lower F1 scores. No solutions for 2000 events and noise rates of 1.0 and 3.0 could be obtained with the homogeneous version of GaMMA. The 1D model provides a solution for 1.0 noise, however with low F1 score and a runtime above one day. We suspect that a key limitation for GaMMA is the DBSCAN algorithm that is used to group the picks before the expectation maximization algorithm. For very dense sequences, this algorithm fails to break up the picks into small enough clusters, and larger clusters take substantially longer to associate.

5 Application to the 2014 Iquique sequence

In addition to the synthetic tests, we evaluate the different associators on a real scenario. For this, we study the 2014 Iquique sequence. Starting with an 8 month long slow slip transient, the 2014 Iquique sequence contained a magnitude 6.6 foreshock on 16th March and the mainshock on the evening of 1st April (Socquet et al., 2017; Soto et al., 2019). We look at the time between 15th March 2014 and 15th April 2014. This time span includes the largest foreshock, the mainshock, and the phase of most intensive aftershock activity. For this study, we use data from the 20 stations in the CX network. We note that generally more stations from other networks are available in the area. However, as we do not aim to produce a comprehensive catalog but rather to test the associators, we restrict ourselves to the high-quality CX stations.

Using the CX data, we build a small earthquake detection workflow. First, we pick P and S arrivals in the continuous waveforms using PhaseNet (Zhu and Beroza, 2019) trained on INSTANCE (Michelini et al., 2021) using SeisBench (Woollam et al., 2022). We use a pick threshold of 0.05 for both P and S waves, i.e., every pick that has a confidence value above 0.05 assigned to it by the deep learning picker is treated as an arrival. This is intentionally a very low threshold to further stress test the associators. Second, we pass the picks to each associator to obtain catalogs. For the homogeneous velocity model, we use 7.0 km/h (P) and 4.0 km/h (S), for the 1D model the one from Graeber and Asch (1999). All remaining parameters are provides in Tables S1, S2 and S3. For each associator, we provide picks in daily chunks. As in our benchmark, we require at least 10 picks and 4 stations with both P and S pick. We note that this is an extremely simplistic catalog generation workflow that misses essential postprocessing steps, such as absolute and relative relocation or magnitude estimation. However, it is sufficient to investigate the difference between the associators.

For this analysis, we compare PyOcto (1D and homogeneous model), REAL (1D and homogeneous model), and GaMMA (homogeneous model). We exclude GaMMA with a 1D model, as the model failed to converge for ∼30% of the days given 24 h compute time per day. We reproduced this behaviour multiple times to rule out stochastic artifacts. No configurations have been changed between GaMMA with a homogeneous model and the 1D model except for the velocity model. A visualisation of the partial catalog obtained using GaMMA with a 1D model is available in Figure S2.

Figure 5 shows the seismicity in the IPOC area, including Northern Chile, as determined with the different associators. All catalogs clearly show the main features of the seismicity: an intense cluster of events around the Iquique mainshock in the North-West, moderate seismicity along the subducting slab, and a strong band of deeper seismicity. Table 3 shows statistics for the number of events per catalog, the number of associated picks and the fraction of total picks associated. Overall, the PyOcto and REAL catalogs are largest, with the catalogs from REAL containing slightly more events. For both PyOcto and REAL, the catalogs with homogeneous velocity models are slightly larger. This is most likely related to the different choice in travel time tolerances. The catalog from GaMMA is about a quarter smaller. Overall, PyOcto and REAL associated between 43 % and 46 % of all picks while GaMMA associated 34 %. We note that this does not imply that all remaining picks are incorrect, as many might stem from events that have not been recorded at sufficiently many stations to meet the quality control criteria or even be associated.

Figure 6 shows the daily number of events and the average number of P and S picks per event per day. Across all days, the number of events is very similar between all variants of REAL and PyOcto, with PyOcto always detecting slightly more events than REAL in the early parts of the aftershock sequence. GaMMA consistently finds fewer events, with the absolute and relative difference

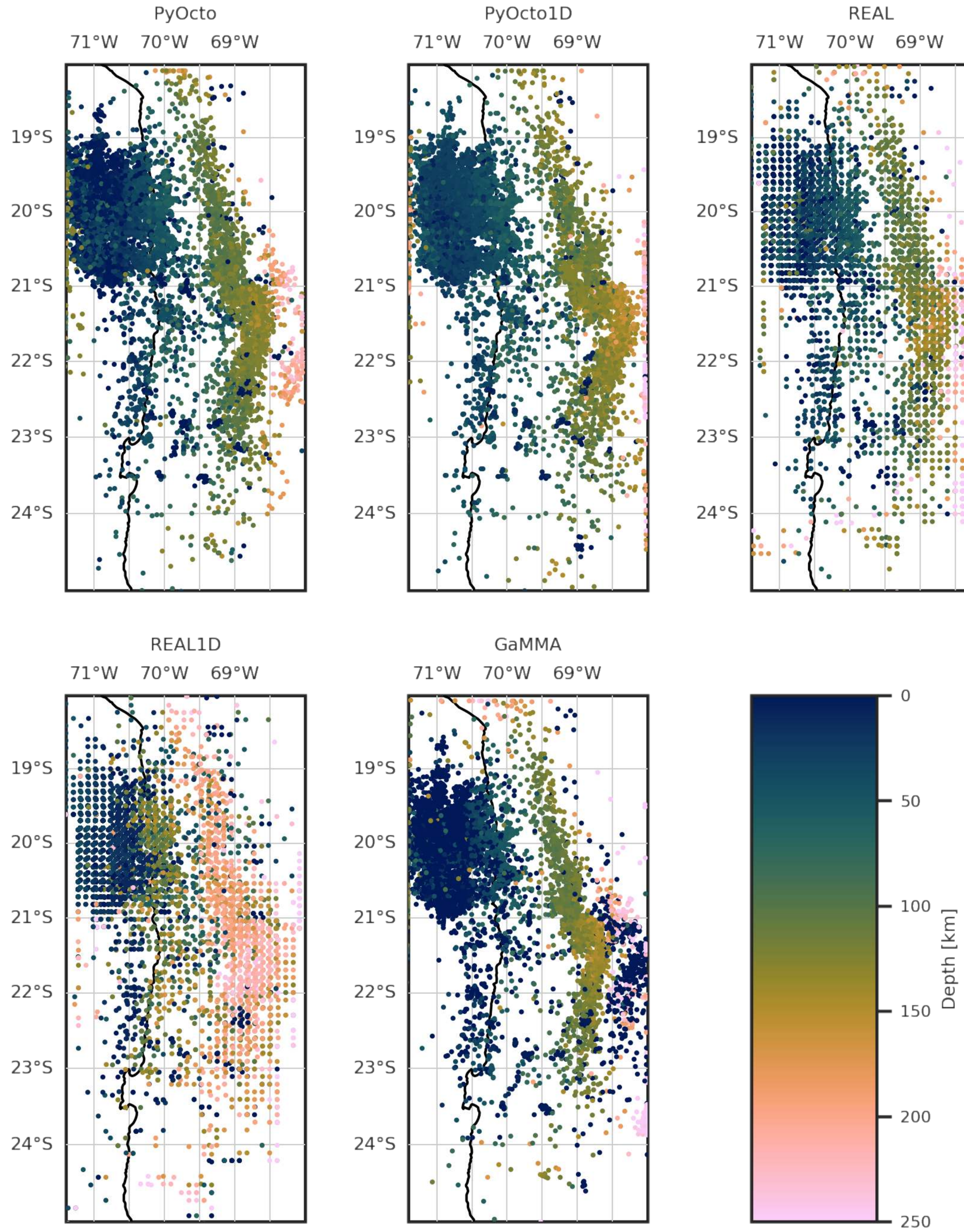


Figure 5 Catalogs generated for the Iquique sequence (15th March 2014 to 15th April 2014) using different phase associators. We visualize the output locations as provided by the associators. Please note that in a comprehensive workflow, absolute and relative relocation techniques should be used as a refinement step. Cross section plots are shown in Figure S3. The station configuration is shown in Figure S1.

becoming particularly large on days with high seismicity rate. This indicates that GaMMA is less able to deal with high rates of seismicity. Notably, for all models the number of detected events stays almost constant for four days after the mainshock. As typically a clear decay in the number of aftershocks in this time frame would be expected, this suggests that all models miss events during these days. Our results can not distinguish if this is a limitation of the picking model or the association models.

Looking at the average number of picks per event, the only noticeable difference between the associators is that REAL consistently finds about 0.6 S picks more per event than PyOcto (Figure 6). Differences in the number of picks are related to the tolerance criteria applied for matching the picks to origins. As the different asso-

Table 3 Catalog statistics for the Iquique sequence catalog with different associators. The table shows the number of events, picks per event, the fraction of associated picks among all picks, and the total number of picks. We abbreviate *picks per event* as *ppe*. Times refer to average run times per day of data.

Associator	Events	Ppe	P ppe	S ppe	Associated	P associated	S associated	Total picks	Time [s]
GaMMA	12,718	16.92	9.90	7.02	0.34	0.31	0.39	634,647	1021
PyOcto	16,660	16.77	9.62	7.15	0.44	0.39	0.52	634,647	12
PyOcto1D	16,362	16.56	9.49	7.06	0.43	0.39	0.50	634,647	15
REAL	16,747	17.35	9.66	7.69	0.46	0.40	0.56	634,647	1487
REAL1D	16,489	17.51	9.78	7.73	0.46	0.40	0.55	634,647	1557

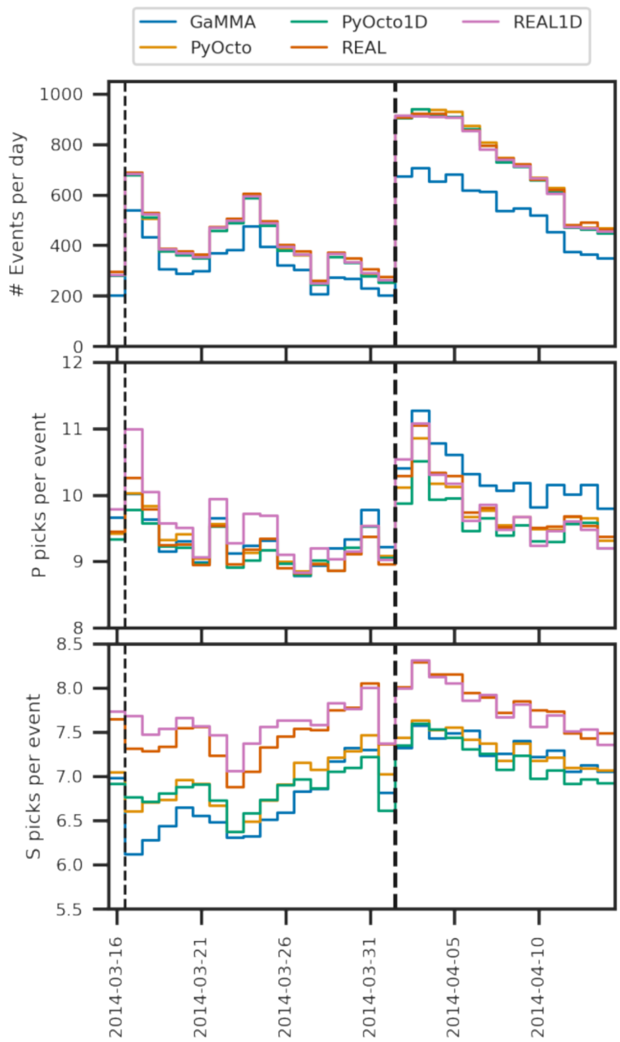


Figure 6 Daily earthquake rates, daily number of associated P picks per event, and daily number of associated S picks for the catalogs generated using the different associators. Vertical black lines indicate the times of the largest foreshock and the mainshock.

ciators use slightly different criteria, it is hard to achieve identical settings. Therefore, the difference in number of picks is likely related to the choice of tolerance parameters. It is difficult to quantify how many of the additional picks are correct or false picks.

An interesting aspect is the temporal development of picks per event. Overall, the number of P picks per event seems to correlate slightly positively with the total number of events. For the S picks, the rate of association also follows systematic patterns across all associators, but a correlation with the number of events is not as apparent. We suggest that the shifts in the number of associated picks are related to the picker performance over time, which is in turn affected by the event distribution. More large events will cause more impulsive, i.e., easier to detect arrivals. At the same time, a higher seismicity rate will also cause higher noise levels, making phase detection and picking overall more challenging.

While this study does not focus on the location accuracy of different associators, as we do not perceive this as the main output of the phase associators, we still provide a brief analysis of our findings in the Iquique sequence. Each method produces a distinct signature of location artifacts (Figures 5 and S3). GaMMA features a substantial number of shallow detections not present in the other catalogs. These are primarily mislocations, likely caused by the initialisation of the sources for the expectation-maximization algorithm at the surface. They occur primarily outside the network. REAL shows clear gridding artifacts caused by the discretisation of the search grid. Finer search-grids would reduce this effect, but come at a substantial compute cost, with halving the grid-space leading to roughly 8 times longer run time. PyOcto shows line-shaped artifacts, however, these are particularly visible with regard to event depth. These stripes are caused by failures in the minimization of the EDT loss in the localization procedure. The artifact is more pronounced for the homogeneous velocity model than the 1D velocity model, likely because the EDT loss is more focused for the 1D model. Stripes could be reduced or eliminated by increasing the sampling depth in the octotree search for localization. However, this would lead to increased runtime. In conclusion, while all associators give a good overview of the general spatial patterns of the seismicity, the locations should only be treated as preliminary estimates. For accurate location, absolute or relative relocation tools, e.g., NonLinLoc (Lomax et al., 2000) or HypoDD (Waldhauser, 2001), should be employed.

We measured average runtimes per day for each associator. As in the synthetic benchmark, PyOcto was by far the fastest, taking 12 s (homogeneous) / 15 s (1D model). Gamma took about 17 minutes per day, REAL took 25 minutes (homogeneous) / 26 minutes (1D model). This means a speed-up factor of 70 to 130 for PyOcto compared to the baselines. As a reference, loading the waveform data from disk and picking it took around

60 to 90 s per day. This means that in this scenario, run times for PyOcto association are one order of magnitude below the times for picking, while for the other associators the association largely dominates the total run time.

To analyse the influence of the minimum required number of picks on the catalog, we conducted additional tests requiring only 7 (instead of 10) total picks per event and only 3 (instead of 4) stations with both P and S pick. The results are shown in Figures S4, S5 and S6. Overall, the results show the same trends as with the more strict requirements. The number of events increases by around 15 % to 30 %, depending on the associator. In this configuration, PyOcto consistently finds more events than both GaMMA and REAL. On the other hand, the seismicity now also appears substantially more scattered for all associators. It is unclear, to which degree this is caused by incorrect associations or by less accurate locations from the lower number of picks. The runtimes of the three associators stay largely unaffected by the change in pick requirements.

6 Conclusion

In this paper, we introduced PyOcto, a novel seismic phase associator based on space-time partitioning. We tested PyOcto in two distinct synthetic earthquake scenarios with different numbers of events and different noise levels. PyOcto consistenly showed detection performance on par or even superior to the state of the art approaches GaMMA and REAL. At the same time, PyOcto achieves substantial speedups, often with factors above 50. We furthermore compared the algorithms on the challenging 2014 Iquique sequence. Here too, PyOcto produces a very complete seismicity catalog. Similar to the synthetic cases, PyOcto again achieves a speedup of above 70 compared to the other associators, with phase association taking substantially shorter time than the phase picking. This makes the algorithm future-proof in face of ever-growing seismic networks and potentially more sensisitive, future phase pickers. PyOcto is available as an open-source tool.

Acknowledgements

This work has been partially supported by MIAI@Grenoble Alpes (ANR-19-P3IA-0003). This work has been partially funded by the European Union under the grant agreement n°101104996 (“DECODE”). Views and opinions expressed are however those of the author only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them. I thank Frederik Tilmann and Marius Isken for insightful discussions that helped improve the algorithm design. I thank Sophie Giffard-Roisin, Stephen Hicks, Sacha Lapins, and an anonymous reviewer for their comments that helped improve the manuscript.

Data and code availability

PyOcto is available at https://github.com/yetinam/pyocto and at https://doi.org/10.5281/zenodo.10016665. The code for the benchmark is available in the same repository. PyOcto can be installed from PyPI using pip. Waveform data for the CX network (https://doi.org/10.14470/PK615318) was obtained through the GEOFON FDSN webservice.

Competing interests

The author has no competing interests.

References

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., and Wassermann, J. ObsPy: A Python toolbox for seismology. *Seismological Research Letters*, 81(3):530–533, 2010. doi: 10.1785/gssrl.81.3.530.

Ester, M., Kriegel, H.-P., Sander, J., Xu, X., et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, volume 96, pages 226–231, 1996. doi: 10.5555/3001460.3001507.

GFZ German Research Centre For Geosciences and Institut Des Sciences De L'Univers-Centre National De La Recherche CNRS-INSU. IPOC Seismic Network, 2006. doi: 10.14470/PK615318.

González-Vidal, D., Moreno, M., Sippl, C., Baez, J. C., Ortega-Culaciati, F., Lange, D., Tilmann, F., Socquet, A., Bolte, J., Hormazabal, J., et al. Relation between oceanic plate structure, patterns of interplate locking and microseismicity in the 1922 Atacama seismic gap. *Geophysical Research Letters*, 50(15): e2023GL103565, 2023. doi: 10.1029/2023GL103565.

Graeber, F. M. and Asch, G. Three-dimensional models of P wave velocity and P-to-S velocity ratio in the southern central Andes by simultaneous inversion of local earthquake data. *Journal of Geophysical Research: Solid Earth*, 104(B9):20237–20256, 1999. doi: 10.1029/1999JB900037.

Johnson, C. E., Bittenbinder, A., Bogaert, B., Dietz, L., and Kohler, W. Earthworm: A flexible approach to seismic network processing. *Iris newsletter*, 14(2):1–4, 1995.

Liu, M., Zhang, M., Zhu, W., Ellsworth, W. L., and Li, H. Rapid characterization of the July 2019 Ridgecrest, California, earthquake sequence from raw seismic data using machine-learning phase picker. *Geophysical Research Letters*, 47(4):e2019GL086189, 2020. doi: 10.1029/2019GL086189.

Lomax, A., Virieux, J., Volant, P., and Berge-Thierry, C. Probabilistic earthquake location in 3D and layered models: Introduction of a Metropolis-Gibbs method and comparison with linear locations. *Advances in seismic event location*, pages 101–134, 2000. doi: 10.1007/978-94-015-9536-0_5.

McBrearty, I. W. and Beroza, G. C. Earthquake phase association with graph neural networks. *Bulletin of the Seismological Society of America*, 113(2):524–547, 2023. doi: 10.1785/0120220182.

Michelini, A., Cianetti, S., Gaviano, S., Giunchi, C., Jozinović, D., and Lauciani, V. INSTANCE–the Italian seismic dataset for machine learning. *Earth System Science Data*, 13(12):5509–5544, 2021. doi: 10.5194/essd-13-5509-2021.

Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., and Beroza, G. C. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature communications*, 11(1):3952, 2020. doi: 10.1038/s41467-020-17591-w.

Moutote, L., Itoh, Y., Lengliné, O., Duputel, Z., and Socquet, A. Evidence of a transient aseismic slip driving the 2017 Valparaiso earthquake sequence, from foreshocks to aftershocks. *Journal of Geophysical Research: Solid Earth*, page e2023JB026603, 2023. doi: 10.1029/2023JB026603.

Münchmeyer, J., Woollam, J., Rietbrock, A., Tilmann, F., Lange, D., Bornstein, T., Diehl, T., Giunchi, C., Haslinger, F., Jozinović, D., et al. Which picker fits my data? A quantitative evaluation of deep learning based seismic pickers. *Journal of Geophysical Research: Solid Earth*, 127(1):e2021JB023499, 2022. doi: 10.1029/2021JB023499.

Ross, Z. E., Yue, Y., Meier, M.-A., Hauksson, E., and Heaton, T. H. PhaseLink: A deep learning approach to seismic phase association. *Journal of Geophysical Research: Solid Earth*, 124(1): 856–869, 2019. doi: 10.1029/2018JB016674.

Ross, Z. E., Zhu, W., and Azizzadenesheli, K. Neural mixture model association of seismic phases. *arXiv preprint arXiv:2301.02597*, 2023. doi: 10.48550/arXiv.2301.02597.

Sippl, C., Schurr, B., Asch, G., and Kummerow, J. Seismicity structure of the northern Chile forearc from> 100,000 double-difference relocated hypocenters. *Journal of Geophysical Research: Solid Earth*, 123(5):4063–4087, 2018. doi: 10.1002/2017JB015384.

Smith, J. D., Azizzadenesheli, K., and Ross, Z. E. Eikonet: Solving the eikonal equation with deep neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 59(12):10685–10696, 2020. doi: 10.1109/TGRS.2020.3039165.

Socquet, A., Valdes, J. P., Jara, J., Cotton, F., Walpersdorf, A., Cotte, N., Specht, S., Ortega-Culaciati, F., Carrizo, D., and Norabuena, E. An 8 month slow slip event triggers progressive nucleation of the 2014 Chile megathrust. *Geophysical Research Letters*, 44(9): 4046–4053, 2017. doi: 10.1002/2017GL073023.

Soto, H., Sippl, C., Schurr, B., Kummerow, J., Asch, G., Tilmann, F., Comte, D., Ruiz, S., and Oncken, O. Probing the northern Chile megathrust with seismicity: The 2014 M8. 1 Iquique earthquake sequence. *Journal of Geophysical Research: Solid Earth*, 124(12): 12935–12954, 2019. doi: 10.1029/2019JB017794.

Tan, Y. J., Waldhauser, F., Ellsworth, W. L., Zhang, M., Zhu, W., Michele, M., Chiaraluce, L., Beroza, G. C., and Segou, M. Machine-learning-based high-resolution earthquake catalog reveals how complex fault structures were activated during the 2016–2017 Central Italy sequence. *The Seismic Record*, 1(1): 11–19, 2021. doi: 10.1785/0320210001.

Waldhauser, F. hypoDD–A program to compute double-difference hypocenter locations. 2001. doi: 10.7916/D8SN072H.

Wilding, J. D., Zhu, W., Ross, Z. E., and Jackson, J. M. The magmatic web beneath Hawai ʻi. *Science*, 379(6631):462–468, 2023. doi: 10.1126/science.ade5755.

Woollam, J., Rietbrock, A., Leitloff, J., and Hinz, S. Hex: Hyperbolic event extractor, a seismic phase associator for highly active seismic regions. *Seismological Research Letters*, 91(5):2769–2778, 2020. doi: 10.1785/0220200037.

Woollam, J., Münchmeyer, J., Tilmann, F., Rietbrock, A., Lange, D., Bornstein, T., Diehl, T., Giunchi, C., Haslinger, F., Jozinović, D., et al. SeisBench—A toolbox for machine learning in seismology. *Seismological Research Letters*, 93(3):1695–1709, 2022. doi: 10.1785/0220210324.

Zhang, M., Ellsworth, W. L., and Beroza, G. C. Rapid earthquake association and location. *Seismological Research Letters*, 90(6): 2276–2284, 2019. doi: 10.1785/0220190052.

Zhu, W. and Beroza, G. C. PhaseNet: A deep-neural-network-based seismic arrival-time picking method. *Geophysical Journal International*, 216(1):261–273, 2019. doi: 10.1093/gji/ggy423.

Zhu, W., McBrearty, I. W., Mousavi, S. M., Ellsworth, W. L., and Beroza, G. C. Earthquake phase association using a Bayesian Gaussian mixture model. *Journal of Geophysical Research: Solid Earth*, 127(5):e2021JB023249, 2022. doi: 10.1029/2021JB023249.

doi:10.26443/seismica.v3i1.1128

SEISMICA

ScS shear-wave splitting in the lowermost mantle: Practical challenges and new global measurements

Jonathan Wolf * [1], **Maureen D Long** [1]

[1]Department of Earth and Planetary Sciences, Yale University, New Haven, CT, USA

Author contributions: *Conceptualization*: Jonathan Wolf, Maureen Long. *Methodology*: Jonathan Wolf. *Software*: Jonathan Wolf. *Formal Analysis*: Jonathan Wolf. *Investigation*: Jonathan Wolf. *Writing - Original draft*: Jonathan Wolf. *Writing - Review & Editing*: Jonathan Wolf, Maureen Long. *Visualization*: Jonathan Wolf. *Supervision*: Maureen Long.

Abstract Many regions of the Earth's mantle are seismically anisotropic, including portions of the lowermost mantle, which may indicate deformation due to convective flow. The splitting of ScS phases, which reflect once off the core-mantle boundary (CMB), is commonly measured to identify lowermost mantle anisotropy, although some challenges exist. Here, we use global wavefield simulations to evaluate commonly used approaches to inferring a lowermost mantle contribution to ScS splitting. We show that due to effects of the CMB reflection, only the epicentral distance range between $60°$ and $70°$ is appropriate for ScS splitting measurements. For this distance range, splitting is diagnostic of deep mantle anisotropy if no upper mantle anisotropy is present; however, if ScS is also split due to upper mantle anisotropy, the reliable diagnosis of deep mantle anisotropy is challenging. Moreover, even in the case of a homogeneously anisotropic deep mantle region sampled from a single azimuth by multiple ScS waves with different source polarizations (in absence of upper mantle anisotropy), different apparent fast directions are produced. We suggest that ScS splitting should only be measured at "null" stations and conduct such an analysis worldwide. Our results indicate that seismic anisotropy is globally widespread in the deep mantle.

Production Editor:
Gareth Funning
Handling Editor:
Lauren Waszek
Copy & Layout Editor:
Hannah F. Mark

Signed reviewer(s):
Sanne Cottaar

Received:
October 14, 2023
Accepted:
April 11, 2024
Published:
April 26, 2024

1 Introduction

Convective flow in Earth can lead to the preferential alignment of minerals, causing waves to travel through the material with different speeds dependent on propagation and polarization directions, a property called seismic anisotropy (e.g., Silver and Chan, 1991; Long and Becker, 2010). Analogous to optical birefringence, shear waves split into a fast and a slow traveling component in seismically anisotropic materials (e.g., Silver and Chan, 1991). Seismic anisotropy has been found to be most prominent in the upper and lower layers of Earth's mantle, while it is almost absent in the bulk of the lower mantle (e.g., Panning and Romanowicz, 2006; Chang et al., 2015). For example, anisotropy has been measured in Earth's crust (e.g., Barruol and Kern, 1996; Haws et al., 2023), the upper mantle (e.g., Silver, 1996; Savage, 1999; Zhu et al., 2020), the mantle transition zone (e.g., Yuan and Beghein, 2014; Chang and Ferreira, 2019) and the uppermost lower mantle (e.g., Foley and Long, 2011; Mohiuddin et al., 2015). Moreover, the lowermost 200-300 km of the mantle, also called D″, is anisotropic in many places (e.g., Kendall and Silver, 1996; Garnero and Lay, 1997; Nowacki et al., 2010; Reiss et al., 2019; Nowacki and Cottaar, 2021; Wolf et al., 2024; see summary by Wolf et al., 2023c).

On average, seismic anisotropy in Earth's upper mantle is stronger than at the base of the mantle (e.g., Panning and Romanowicz, 2006; French and Romanowicz, 2014). It is thus challenging to measure seismic anisotropy in the lowermost mantle because the potential contribution of upper mantle anisotropy to every seismogram needs to be accounted for, as the seismic waves used to infer D″ anisotropy travel through the upper as well as the deepest mantle (e.g., Wolf et al., 2022b). To account for the upper mantle contribution, multiple techniques have been developed, most of which rely on comparisons of the shear wave splitting contribution to multiple seismic waves. A popular method to infer deep mantle anisotropy is from differential splitting of the SKS and SKKS phase (e.g., Wang and Wen, 2004; Niu and Perez, 2004; Long, 2009; Reiss et al., 2019; Wolf et al., 2024). SKS and SKKS have very similar raypaths through the upper mantle and a much larger spatial raypath separation in the lowermost mantle. Therefore, large differences in SKS and SKKS splitting for the same source-receiver pair must be due to lowermost mantle anisotropy (e.g., Niu and Perez, 2004; Wang and Wen, 2004). Alternatively, the splitting of SKS and S_{diff} can be compared. If SKS is not influenced by seismic anisotropy but S_{diff} clearly is, this is evidence for deep mantle anisotropy causing splitting of S_{diff} (Cottaar and Romanowicz, 2013; Wolf et al., 2023b; Wolf and Long, 2023). The advantage of measurements using SKS, SKKS and S_{diff} waves is that the source-side anisotropy contribution in the upper mantle is either erased by the P-to-SV conversion at the core-mantle boundary (CMB; SKS and SKKS) or, under certain conditions, negligible (S_{diff};

*Corresponding author: jonathan.wolf@yale.edu

Wolf et al., 2023b). A technique that explicitly accounts for source-side and receiver-side anisotropy is S-ScS differential splitting (Wookey et al., 2005). After applying explicit ray-theoretical corrections to S and ScS for known receiver-side anisotropy, source-side anisotropy can be measured from the corrected S phase. After correcting ScS for the source-side contribution, the remaining anisotropy contribution to ScS must be due to D″ anisotropy (Wookey et al., 2005b; Nowacki et al., 2010; Creasy et al., 2017; Pisconti et al., 2023).

These differential splitting techniques make a number of assumptions, typically in the context of ray theory. These assumptions have been tested using global wavefield simulations. For example, the interpretation of differential SV_{diff}-SH_{diff} travel times as being uniquely indicative of D″ anisotropy has been questioned (Komatitsch et al., 2010; Borgeaud et al., 2016; Parisi et al., 2018) as isotropic models can induce SV_{diff}-SH_{diff} travel time differences under certain circumstances. The SKS-SKKS differential splitting technique, on the other hand, has largely been shown to reliably detect anisotropy if certain caveats are considered (Tesoniero et al., 2020; Wolf et al., 2022b; see also, Lin et al., 2014). Nowacki and Wookey (2016) pointed out that some of the ray-theoretical assumptions do not always hold for the S-ScS differential splitting technique, especially in case of heterogeneous anisotropy. In particular the assumption of a horizontal ScS raypath through D″ is a significant oversimplification. Additionally, Wolf et al. (2022b) showed that the phase shift of the radial component of ScS due to the reflection off the mantle-core interface needs to be explicitly considered to accurately measure ScS splitting. Also, Parisi et al. (2018) demonstrated that differential ScS SV-SH travel times can be produced by isotropic structure at distances $> 90°$. Some of these challenges could successfully be resolved; for example, the horizontal raypath assumption has been avoided in recent S-ScS differential splitting studies (e.g., Pisconti et al., 2023; Asplet et al., 2023). However, there still are many open questions, the answers to which will help our ability to use ScS to measure deep mantle anisotropy.

In this work, we assess in detail how ScS waves can be used to measure D″ anisotropy. To do so, we address several questions. First, we analyze the effects of the CMB reflection on the polarization of ScS, and how they influence the measured ScS splitting parameters. Second, we use global wavefield simulations to investigate whether and how apparent shear-wave splitting can be produced for isotropic input models. Here we use the term ‘shear-wave splitting measurements’ to refer to the measurement of splitting parameters (delay time, fast polarization direction, splitting intensity) and not simply to differential SV-SH delay times. (This distinction is important, because shear-wave splitting defined in this way includes requirements regarding the waveform’s shape.) Third, we analyze how well the source-side correction of the S-ScS splitting technique works in light of the polarization effects on ScS due to its CMB reflection and the slightly different raypaths of S and ScS in the source-side upper mantle. Fourth, we assess the accuracy of explicit ScS receiver-side anisotropy corrections using a realistic forward modeling framework. Putting all these insights together, we suggest a strategy for inferring deep mantle anisotropy from the shear wave splitting of ScS waves. Fifth, we apply this strategy globally to analyze deep mantle anisotropy using suitable broadband seismic stations. We find evidence for seismic anisotropy in regions that have been analyzed in previous studies, such as beneath the northern Pacific Ocean, the Caribbean and northern Asia, but we also identify deep mantle anisotropy in previously unexplored regions such as beneath southern Russia and the southwestern Pacific Ocean. Finally, we discuss ways forward to improve the reliability of ScS splitting measurements and interpretations for D″ anisotropy studies.

2 Methods

2.1 Global wavefield simulations

We use the global wavefield modeling code AxiSEM3D (Leng et al., 2016; 2019) in this work. While the code can handle arbitrary three dimensional input models, it calculates synthetic seismograms very efficiently in axisymmetric input models, at the same speed as the older AxiSEM code (Nissen-Meyer et al., 2014). We mostly conduct simulations using axisymmetric models such as isotropic PREM (Dziewonski and Anderson, 1981), which we always use as background model, following our previous work (e.g., Wolf et al., 2022a). We always consider PREM-attenuation and Earth’s ellipticity in our simulations. In some simulations (see below), we replace PREM’s mantle velocity structure with the tomographic model S40RTS (Ritsema et al., 2011). We carry out numerical experiments with and without seismic anisotropy in the lowermost and/or upper mantle. In all simulations presented in this work, we compute synthetic seismograms down to minimum periods of 5 s.

Our source-receiver configuration is shown in Figure 1a. We place a strike-slip earthquake at the north pole and stations at epicentral distances of $60°$ to $100°$ (spaced in $1°$ distance increments) between longitudes $0°$ to $90°$ (spaced in $10°$ increments). We choose event depths of either 100 m or 500 km. A strike-slip focal mechanism is selected such that the initial source polarization of S and ScS is purely SH for longitudes $0°$ and $90°$ and purely SV for longitude $45°$ (Figure 1). For each candidate event depth, we conduct three types of simulations:

1. Isotropic simulations:
   (a) using isotropic PREM (Figure 1b) as input model;
   (b) incorporating modified velocities in the lowermost 150 km of the mantle, replacing those of PREM (Figure 1b);
   (c) incorporating a 3D tomography model (S40RTS) in the mantle, replacing PREM velocities.
2. Anisotropic simulations with lowermost mantle anisotropy:

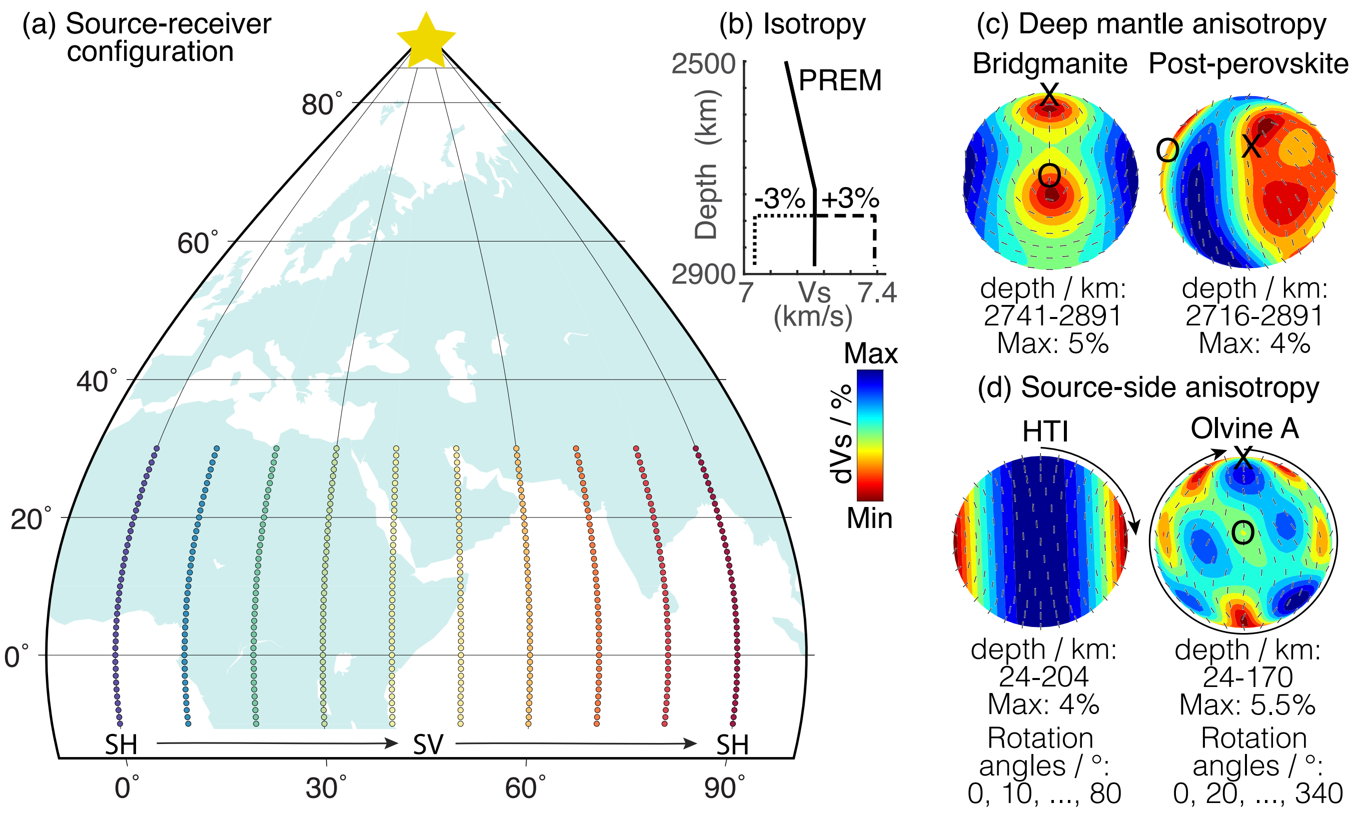


Figure 1 Source-receiver configuration and input models of synthetic simulations. (a) Source and receivers: The strike-slip event (see text) is shown as a yellow star; stations are represented as dots, colored by longitude (which corresponds to the initial polarization of the arriving wave). (b) Lowermost mantle velocity as a function of depth for the PREM model, PREM with 3% reduced (dotted line), and 3% increased velocities (dashed line). All these scenarios are used in our synthetic modeling. (c) Upper hemisphere representations of the elastic tensors (bridgmanite, post-perovskite) used in simulations in which we incorporate deep mantle anisotropy (at the depths shown below each elastic tensor plot). The elastic tensors were taken from the elastic tensor library of Creasy et al. (2020). The color scale shows the percentage of S-wave anisotropy as a function of direction. The maximum percentage is shown at the bottom and depends on the elastic tensor. The small black sticks indicate the fast polarization direction of the S wave for the corresponding propagation direction. The black 'O' represents the shear-plane normal and 'X' the shear direction. The lowermost mantle elastic tensors are oriented such that robust shear-wave splitting measurements can be obtained. (d) Similar to panel (c), for upper mantle source-side anisotropy. The elastic tensor rotation performed in this work is indicated by arrows. The HTI elastic tensor was calculated using MSAT (Walker and Wookey, 2012) and the olivine type-A elastic tensor was taken from Karato (2008).

(a) incorporating bridgmanite (Br) anisotropy in the lowermost 150 km of the mantle;

(b) incorporating post-perovskite (Ppv) anisotropy in the lowermost 175 km of the mantle.

These elastic tensors were taken from Creasy et al. (2020) and are displayed as upper hemisphere representations in Figure 1c. The use of these elastic tensors leads to slightly different lowermost mantle velocities than PREM. The main goal of these simulations is to evaluate the influence of realistic lowermost mantle anisotropy on ScS seismic waves; the isotropic effects are analyzed in the previous set of simulations.

3. Anisotropic simulations with upper mantle anisotropy:

(a) using horizontal transverse isotropy (HTI) in the upper mantle (Figure 1d). The HTI elastic tensor is calculated using MSAT (Walker and Wookey, 2012) and incorporated at the depth range of 24 km to 204 km.

(b) using olivine (A-type fabric) anisotropy in the upper mantle (Figure 1d). The elastic tensor is from Karato (2008) and the anisotropy is incorporated at the depth range of 24 km to 170 km.

In both cases, the anisotropy in the upper mantle leads to a maximum delay time of $\sim$1.5 s. In order to sample anisotropy from different directions, the elastic tensors are rotated around the vertical axis (with respect to their representations in Figure 1d) for different simulations. Due to its symmetry, the HTI elastic tensor is only rotated by angles of $0°$ to $80°$ (in $10°$ increments), while the olivine elastic tensor is rotated between $0°$ to $340°$ (in $20°$ increments).

Synthetic radial and transverse seismograms as a function of distance (for PREM as input model), aligned

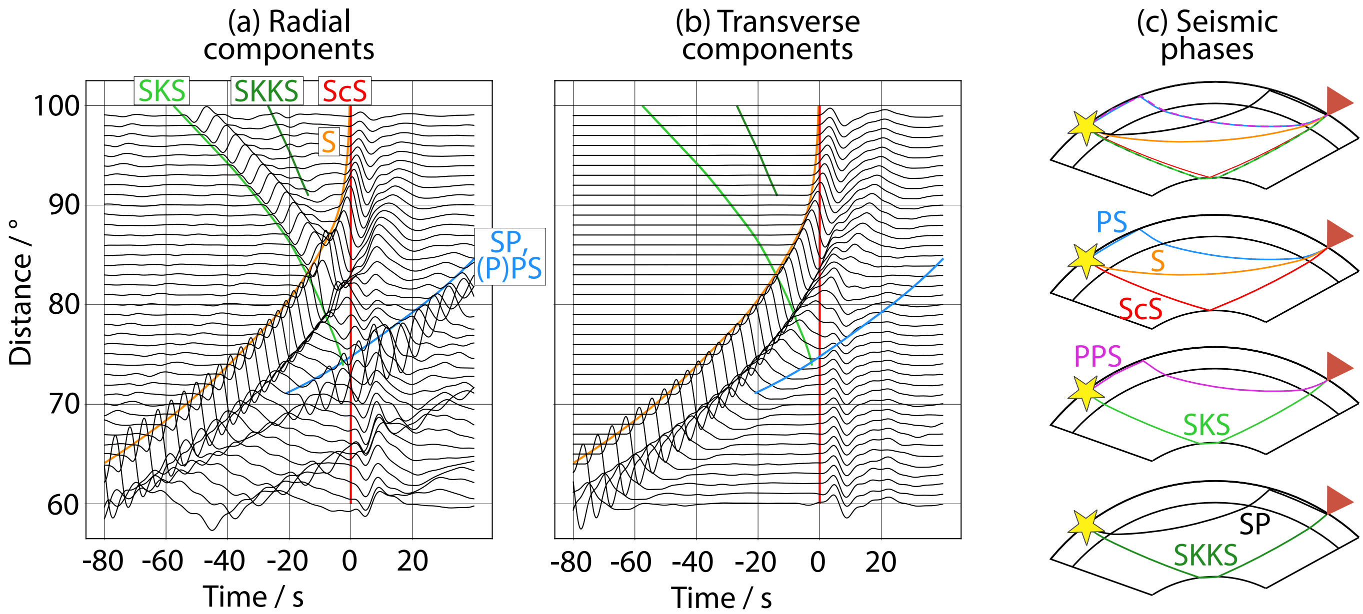


Figure 2 Synthetic waveforms as a function of distance for stations placed along longitude $70°$ (Figure 1), with sketch of relevant seismic phases. (a) Radial component displacement waveforms, plotted at every $1°$ distance increment. Incoming high-amplitude seismic phases are marked with colored lines. (b) Same as panel (a) for the transverse component. (c) Schematic diagrams of raypaths through Earth for the seismic phases marked in panels (a) and (b). The source is shown as a yellow star and the station, at an epicentral distance of $70°$, as a red triangle.

on the predicted ScS arrival, are shown in Figure 2a,b. At an epicentral distance of around $75°$, interference from the PS and PPS phases, which arrive very close together in time at these distances, can be observed. Additionally, some SP energy (which arrives contemporaneously to PS for a $0\,\mathrm{km}$ deep source) likely arrives on the radial component. While PS interference can be observed in the record section shown in Figure 2, the phase is not observable at this distance range for events with focal depths deeper than $200\,\mathrm{km}$, although some PPS energy may still be relevant. For distances $> 80°$, ScS starts to merge with S. For distances $< 70°$ and $> 63°$, SKS and ScS arrive almost contemporaneously, although it is unclear whether SKS has a sufficiently large amplitude to noticeably influence ScS. (A partial answer to this question will be discussed in Section 4.) The raypaths of the seismic phases that may potentially interfere with ScS are shown in a cross-section in Figure 2c.

2.2 Shear wave splitting measurements

Shear wave splitting, which is analogous to optical birefringence, is a consequence of seismic anisotropy. A shear wave that travels through an anisotropic medium splits into a fast and a slow component. The time lag between these components is called δt and the polarization direction of the fast traveling wave is usually referred to as ϕ when measured clockwise from the north, or ϕ' (Nowacki et al., 2010) when measured with respect the incoming wave's backazimuth. Another quantity that is frequently used is the splitting intensity (Chevrot, 2000), SI, which yields a scalar value indicating the splitting strength on an individual seismogram. The splitting intensity is defined as:

$$SI = -2\frac{Pol_{90}(t)Pol_0'(t)}{|Pol_0'(t)|^2} \approx \delta t \sin(-2\phi')\,, \qquad (1)$$

with $Pol'_{0(t)}$ denoting the time derivative of the component in the direction of initial polarization, whereas $Pol_{90}(t)$ is the horizontal seismogram component oriented $90°$ away from the incoming wave's primary polarization.

We determine the splitting parameters (ϕ, δt) using a modified version of the SplitRacer software (Reiss and Rümpker, 2017), which is the same version previously used by Wolf et al. (2022b). This version estimates the initial polarization of the incoming wave, through particle motion analysis, as ScS waves are not typically initially SV-polarized. SplitRacer calculates the splitting parameters (ϕ, δt) using the transverse energy minimization technique (Silver and Chan, 1991), incorporating a corrected calculation of the 95% confidence intervals (Walsh et al., 2013). Whenever we apply source-side anisotropy corrections for the S-ScS differential splitting technique, we measure source-side anisotropy splitting parameters with SplitRacer. Then, we use a code to correct the ScS phase for these source-side splitting parameters, following the algorithm described in Wolf et al. (2022b), which is based on work from Wookey et al. (2005). In this algorithm, we also calculate splitting parameters using the transverse energy minimization technique, building upon an implementation from Creasy et al. (2017). For all these measurements, we consider (ϕ, δt) measurements well-constrained if the 95% confidence intervals are smaller than $\pm 25°$ for ϕ, $\pm 0.8\,\mathrm{s}$ for δt and ± 0.5 for SI. Before measuring splitting parameters, we filter our seismograms retaining periods between $5\,\mathrm{s}$ and $15\,\mathrm{s}$ (unless specified differently).

Polarizations are determined from the seismograms at longer periods (8-25 s) from the long axis of the particle motion ellipse.

3 SV reflection coefficients of ScS at the CMB

In order to understand the potential effects of the CMB reflection on ScS phases, we solve the equations of Chapman (2004) to calculate SV reflection coefficients of ScS at the CMB for PREM velocity structure in the whole mantle, as well as for $3\,\%$ reduced and increased velocities with respect to PREM in the lowermost $150\,\mathrm{km}$ of the mantle (Figure 3). Such velocity variations are realistic for Earth's faster and slower lowermost mantle regions (e.g., Ritsema et al., 2011). We also explore variations of the reflection coefficients as a function of source depth and do not find any substantial differences compared to the $0\,\mathrm{km}$ case shown in Figure 3. We do not compute SH reflection coefficients as the shear wave velocity in the outer core is zero and SH does not couple with P; therefore, all SH energy will be reflected without a phase or amplitude change. Several observations can be made from Figure 3:

- For distances $< \sim 60°$, SV amplitudes are strongly reduced after the reflection. For example, at an epicentral distance of $\sim 30°$ SV loses $\sim 70\,\%$ of its amplitude. This pattern depends on the lowermost mantle velocity and is therefore only possible to account for exactly if the velocity structure at the reflection point is well known. While the SV amplitude effects are complicated, for most distances $< \sim 60°$ the SV phase shift is simple and close to $180°$ (Figure 3).

- For epicentral distances $< 10°$, SV will simply undergo a sign-flip with amplitudes almost unaffected by the reflection.

- For epicentral distances $> \sim 60°$, SV amplitudes are largely unchanged by the reflection and the SV phase shift is between $160°$ and $180°$, depending on distance and deep mantle velocity structure (Figure 3). Because of this, Wolf et al. (2022b) suggested that the description of SV behavior at distances $> \sim 60°$ as a simple sign-flip is accurate enough for the purpose of ScS splitting measurements.

Our analysis of distance-dependent SV reflection coefficients for ScS shows that it is difficult to infer the presence of deep mantle anisotropy for ScS waves at epicentral distances $< \sim 60°$. For these epicentral distances, relative SV-SH amplitudes will be strongly influenced by the deep mantle velocity structure of the region under study, which needs to be precisely accounted for. However, this appears challenging, as the deep mantle velocity structure in any particular deep mantle region is often poorly known. We therefore focus our following analysis on epicentral distances $> 60°$, which is the most frequently used distance range. For example, the S-ScS differential splitting technique has been suggested to be applicable at a distance range between $60°$ and $85°$ (Wookey et al., 2005). There are also multiple previous studies that have analyzed the behavior of S and ScS waves at distances $> \sim 85°$ (e.g., Kendall and Silver, 1996; Pulliam and Sen, 1998; Fouch et al., 2001) to infer deep mantle anisotropy.

4 S and ScS polarizations in isotropic input models

Next, we analyze S and ScS polarizations at epicentral distances between $60°$ and $100°$ using global wavefield simulations. We conduct synthetic simulations for PREM velocity structure in the whole mantle as well as for $3\,\%$ increased and reduced velocities above the CMB (Figure 1a). In Figure 4, we show measured S polarizations for different initial source polarizations of the wave and source depths of $100\,\mathrm{m}$ and $500\,\mathrm{km}$. The results are only weakly influenced by the lowermost mantle velocity, but do depend on source depth for distances $> 90°$. Figure 2 shows that the S wave polarizations are relatively unaffected by interference from other seismic phases at distances $< 80°$, but start to be influenced by ScS at greater distances. Accordingly, measured S polarizations agree very well with the initial source polarizations for distances $< 80°$ (Figure 4). For larger distances, S polarizations are influenced by ScS but still largely agree with the initial source polarization (Figure 4).

For ScS, the pattern of measured polarizations as a function of distance is more complicated (Figure 3). At epicentral distances between $60°$ and $70°$, ScS initial polarizations are approximately opposite the S wave polarization as controlled by the source (Figure 5) due to the approximate SV sign-flip (Figure 3). However, because the sign-flip of SV is not exact (Figure 3), and because of the potential interference with SKS in some of the epicentral distance range (Figure 2), this pattern is by no means perfect. These two effects are hard to distinguish; however, analyzing them in isolation is not required to understand the conditions under which ScS can be used for analyses of lowermost mantle anisotropy, which is the main goal of this study. For distances between $73°$ and $79°$, interference with PS can lead to estimated polarizations close to SV (Figure 5a). For deep sources (Figure 5b), no PS energy arrives; however, PPS and SP may still influence ScS waveforms around this distance range. Exceptions are observed at the stations at azimuths for which the initial polarization is purely SH, as the (P)PS amplitude is zero for them (Figure 5). For distances $> 80°$, S and ScS merge, with S dominating, leading to polarizations that are close to the S initial source polarization (Figure 5). These overall patterns hold for all the different lowermost mantle velocities that we tested (Figure 5).

Considering these results, it appears challenging to measure deep mantle anisotropy reliably from ScS for distances at which the PPS or PS phase potentially interferes with ScS. This corresponds to a distance range between $73°$ and $79°$ for shallow events (e.g., Figure 5) and to distances down as close as $70°$ for an event depth of $150\,\mathrm{km}$. For events deeper than $\sim 200\,\mathrm{km}$ no PS phase arrives at these distances, but PPS may have an influence

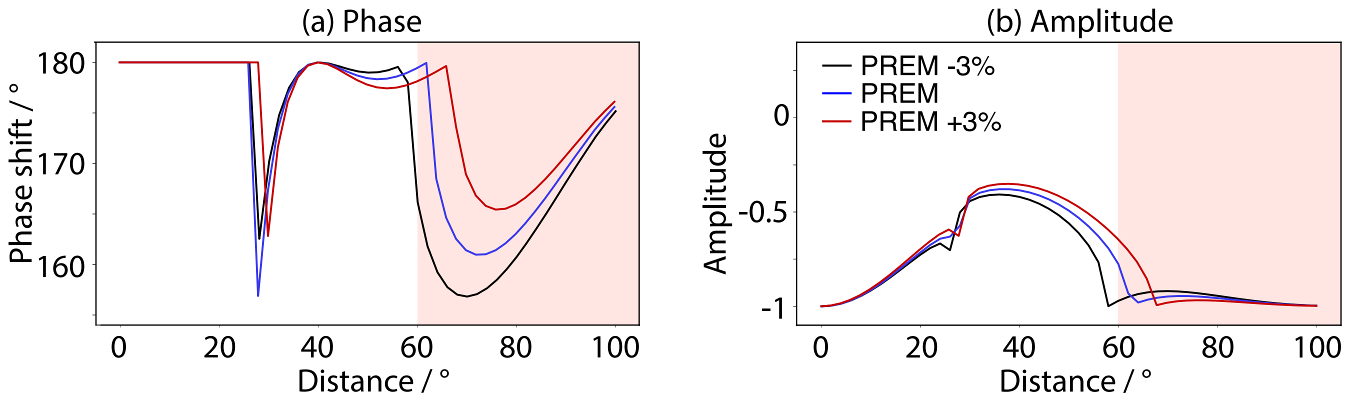


Figure 3 Influence of the CMB reflection on the SV portion of the ScS phase. We calculate reflection coefficients for PREM (blue line), PREM -3% (black line), and PREM +3% (red line) shear velocities in the lowermost $150\,\mathrm{km}$ of the mantle (Figure 1b). The source depth for this calculation is $0\,\mathrm{km}$, but we note that no significant variation can be observed as a function of source depth. Red shading marks epicentral distances $> 60°$, which are typically used for measurements of shear-wave splitting due to lowermost mantle anisotropy. (a) Phase shift of the radial component of the ScS as a function of epicentral distance. (b) Amplitudes (real part of the reflection coefficient) as a function of epicentral distance.

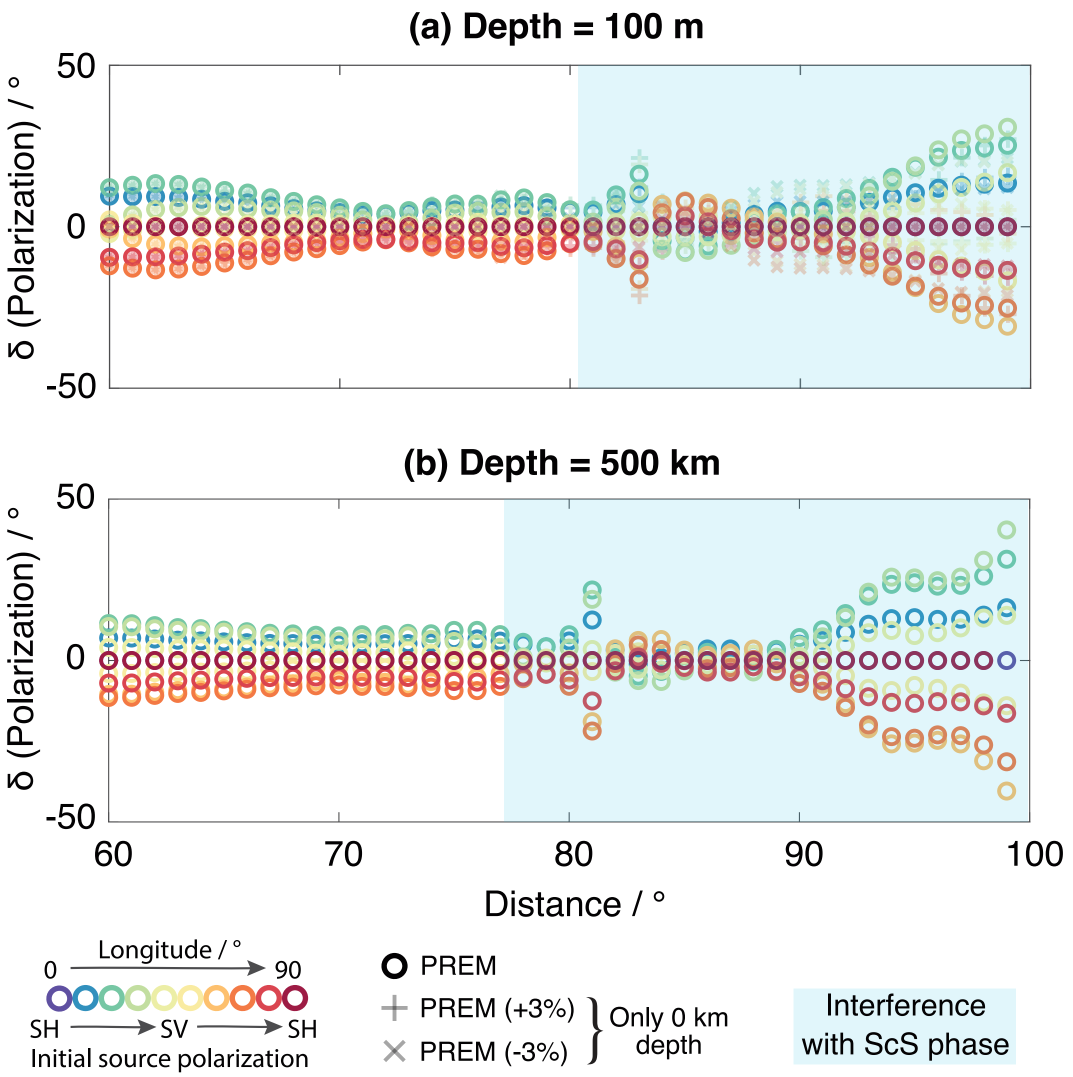


Figure 4 Measured S polarizations relative to the initial source polarization as a function of epicentral distance, for different deep mantle velocity profiles (legend), initial source polarizations (legend), and source depths. Potential interference with ScS is indicated by blue shading. Measurements are conducted for source depths of (a) $100\,\mathrm{m}$ and (b) $500\,\mathrm{km}$.

on seismic waveforms down to $70°$ epicentral distance. Similarly, if shear wave splitting is measured from S/ScS for distances $\sim > 80°$, it should be considered that the S initial polarization likely dominates, but ScS influences

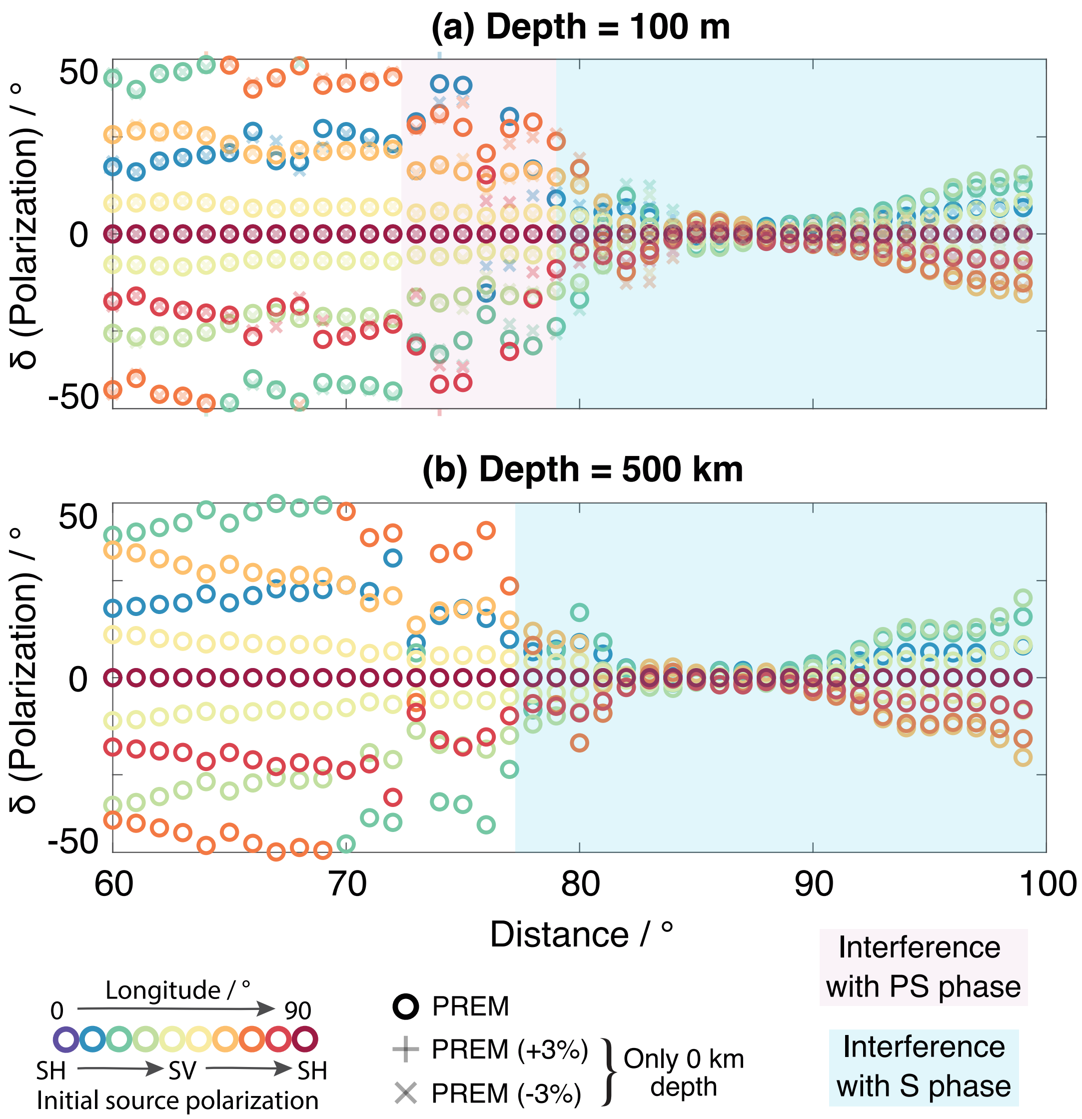


Figure 5 Same as Figure 4 for ScS. Potential PS interference is indicated by pink shading and potential S interference by blue shading.

the waveforms and horizontal component amplitude ratios.

5 Apparent shear-wave splitting in isotropic input models

We have shown how the CMB reflection and phase interferences can influence the polarization of ScS. However, it remains unclear whether such effects can result in apparent shear wave splitting. To test this, we conduct synthetic simulations in isotropic input models as introduced in Section 2.1.

For PREM synthetics, calculated for a focal depth of $100\,\mathrm{m}$, non-null splitting intensities can be reproduced, although we do not measure well-constrained splitting parameters (ϕ, δt) (Figure 6a). If the source is placed in a depth of $500\,\mathrm{km}$, however, apparently well-constrained (ϕ, δt) values can be measured at distances between $90°$ and $94°$ (Figure 6b). Some of these measurements may be identified as null splitting, but not all of them. For PREM+S40RTS, on the other hand, apparently well-constrained (ϕ, δt) values are mainly obtained for distances $> 94°$ (although there is also some apparent splitting at smaller distances). The reason for the apparent splitting is phase interference; for example, the interaction between S and ScS (Figure 2), which arrive at approximately the same time for distances $> 90°$. The transverse components of S and ScS are generally very similar at these distances, as the transverse ScS component is largely unaffected by the reflection. However, the radial ScS component will be approximately sign-flipped compared to S (Figure 3) and potentially have a slightly different amplitude; the details of the phase's behavior depend on lowermost mantle velocity structure and the event depth (Figure 3). If these waveform distortions affect transverse and radial components in a way that the energy on the Pol component has the shape of the time derivative of the Pol_{90} component, apparent splitting results.

We show an example of apparent shear-wave splitting from simulations using isotropic PREM+S40RTS with a source at $100\,\mathrm{m}$ depth in Figure 7. The Pol_{90} component has approximately the shape of the Pol_0 component time derivative (Figure 7a) and the particle motion

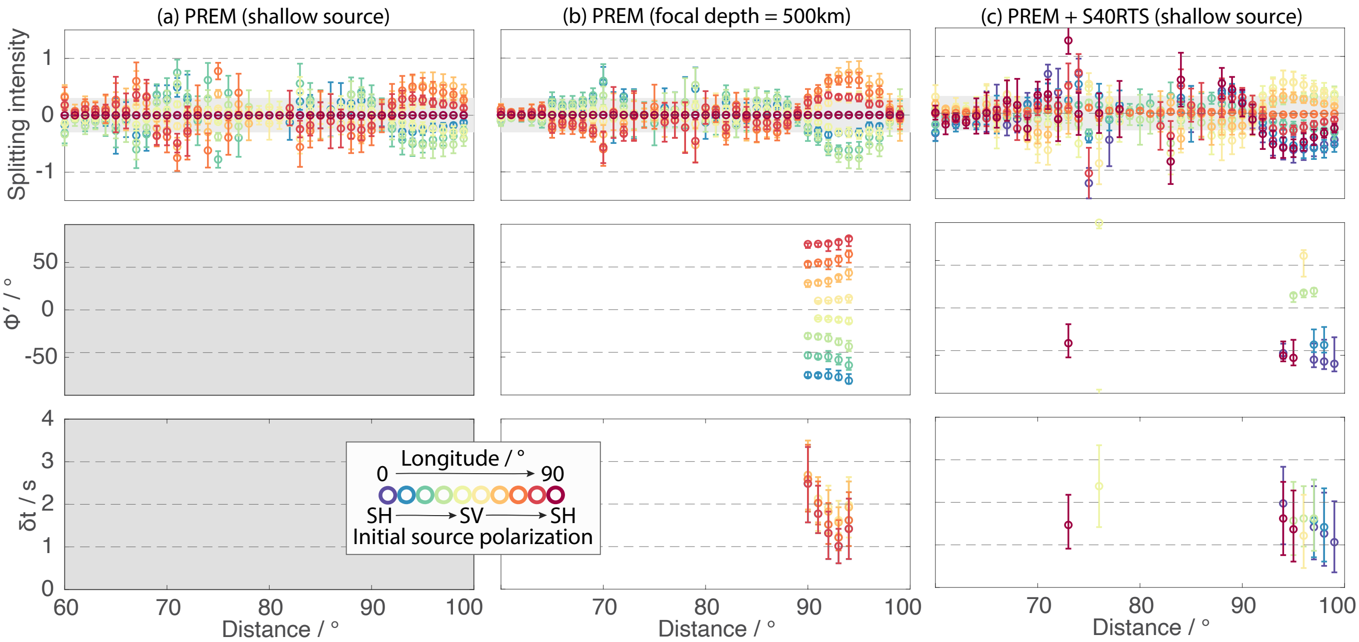


Figure 6 Shear wave splitting parameters SI (top), ϕ' (middle), δt (bottom) as a function of distance, dependent on initial source polarization (legend), for isotropic input models. Error bars indicate $95\,\%$ confidence intervals of the measurements (see Section 2.2). Only well-constrained splitting measurements are shown (see text). (a) For PREM input model and a focal depth of $100\,\mathrm{m}$. (b) For PREM input model and a focal depth of $500\,\mathrm{km}$. (c) For PREM input model, for which the mantle was replaced with the tomographic model S40RTS, and a focal depth of $100\,\mathrm{m}$. Apparent shear-wave splitting can be produced at distances $> 90°$, but also sometimes at smaller distances.

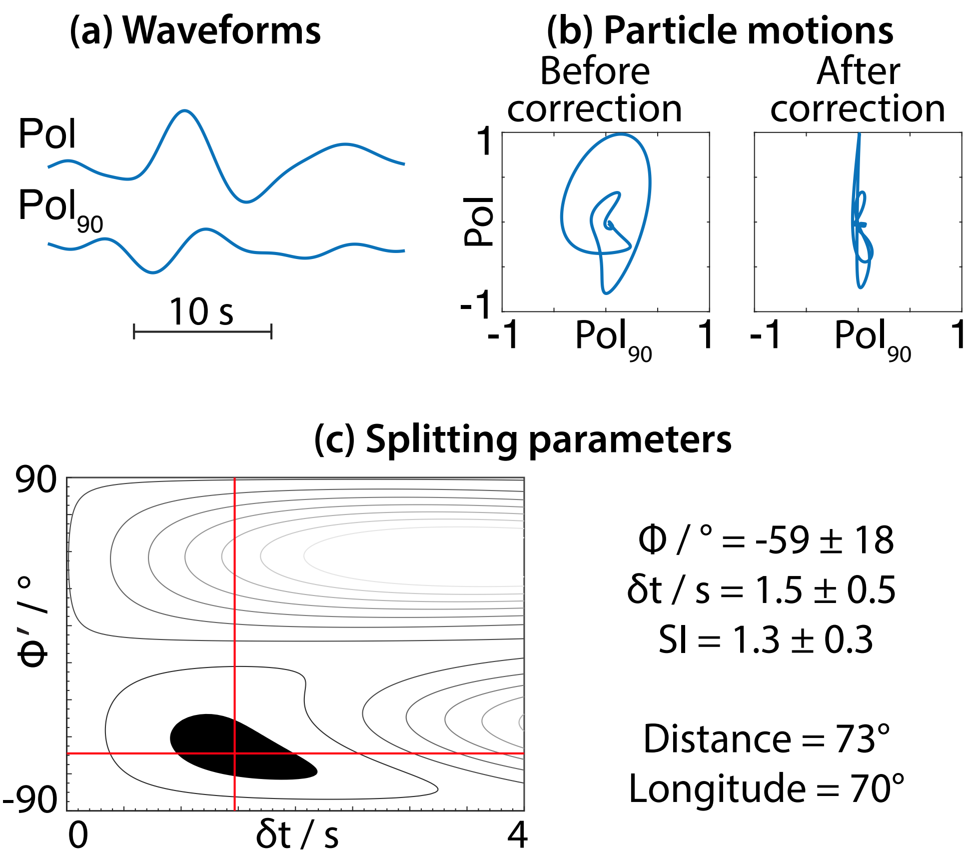


Figure 7 Example apparent shear wave splitting for the (isotropic) PREM+S40RTS input model and a source depth of $0\,\mathrm{km}$. (a) Pol and Pol_{90} waveforms for the combined ScS phase. (b) Particle motions before (left) and after (right) correction for the best-fitting splitting parameters. (c) Best-fitting splitting parameters in the ϕ'-δt plane. The 95% confidence interval is shown in black.

looks elliptical (Figure 7b), mimicking shear wave splitting due to seismic anisotropy. Accordingly, the apparent estimated splitting parameters are well-constrained (Figure 7c).

Our results so far indicate that measurements of shear-wave splitting for epicentral distances $< 60°$ need to carefully consider the SV reflection coefficient at the CMB for ScS, which will depend on the deep mantle velocity structure of the region under study. Additionally, distances between $70°$ and $80°$ cannot be used for ScS splitting measurements if (P)PS or SP may be interfering. For distances $> 80°$, S and ScS merge (Figure 2), making it challenging to distinguish between these phases in seismograms. Apparent splitting of the combined S/ScS phase can be produced in isotropic structure (Figures 6 and 7). Therefore, the most promising distance range to measure ScS splitting due to deep mantle anisotropy is between $60°$ and $70°$.

6 Shear wave splitting in models that incorporate deep mantle anisotropy

We next test the effects of deep mantle anisotropy on measured ScS splitting in absence of upper mantle anisotropy, incorporating Br (Figure 8a) and Ppv (Figure 8b) anisotropy in the lowermost mantle, replacing PREM velocity structure (see Section 2.1). In Figure 8 we show measured shear wave splitting parameters (SI; ϕ', δt) from ScS as a function of epicentral distance and initial source polarization. Due to the aforementioned challenges at many epicentral distances, we focus on shear wave splitting measured at distances between $60°$ and $70°$. In this distance range, we measure many well-constrained (ϕ', δt) values for our anisotropic input models (Figure 8). The seismic anisotropy in the input model is incorporated such that it is sampled from the same direction independent of azimuth. However, we can observe a large spread of measured ϕ' values for both elastic tensors we tested. The reason for this is that the measured splitting is a combination of the splitting accumulated on both legs of the ScS raypath through D$''$ (Figure 8c). The initial polarization of ScS depends on its azimuth in our simulations, and this initial polarization affects how the wave is split on both legs of the raypath. This situation is analogous to splitting from multiple layers of anisotropy in the upper mantle, which produces apparent splitting that depends on azimuth (Silver and Savage, 1994; Silver and Long, 2011). Therefore, it is logical that measured fast polarizations are not the same, even though the same deep mantle anisotropy is sampled. As a consequence, if ScS splitting due to D$''$ anisotropy is measured from a certain sampling direction for any given lowermost mantle region, ScS splitting parameters (ϕ, δt) cannot be expected to be the same for different events, unless the events all have similar initial polarizations. Therefore, the mean splitting measurement as often determined in ScS splitting studies (e.g., Nowacki et al., 2010; Wolf et al., 2019; Pisconti et al., 2023) does not have a clear meaning for the interpretation of mantle flow directions, since the same measurement can be obtained for a variety of anisotropy scenarios and initial polarizations of the wave.

7 Correction of ScS for source-side anisotropy contribution inferred from S

The S-ScS differential splitting technique isolates the lowermost mantle anisotropy contribution to ScS by correcting the ScS waveform for the influence of receiver-side and source-side anisotropy in the upper mantle (Wookey et al., 2005). The source-side anisotropy contribution is inferred from the S waveform, which has been first corrected for the influence of receiver-side anisotropy. The assumptions made in this process are that S and ScS raypaths through the upper mantle are sufficiently similar that both phases experience the same splitting due to upper mantle anisotropy, and that their initial source polarizations are also similar. In the most extreme case, for a source-receiver distance of $60°$ and a surface event, S and ScS raypaths are up to $250\,\mathrm{km}$ apart at the bottom of the transition zone, so that the assumption that S and ScS raypaths are sufficiently close together may only be valid in cases of relatively generally homogeneous upper mantle anisotropy. To account for the CMB reflection, Wolf et al. (2022b) suggested approximating the phase shift of ScS$_{SV}$ as a simple sign-flip of the radial component. More accurate corrections would be challenging, given that the precise phase shift depends on the deep mantle velocity structure near the ScS reflection point (Figure 3). Additionally, our results for a distance range close to $60°$, at which the PREM-predicted phase shift corresponds to a precise sign-flip (Figure 3), do not indicate that ScS splitting measurements could be substantially improved by implementing the PREM-predicted phase shift. Using this assumption, Wolf et al. (2022b) showed that approximate source-side splitting parameters for ScS can indeed be inferred from S. These splitting parameters can then be used to correct ScS waveforms after a correction for receiver side anisotropy has been applied (Wookey et al., 2005).

Measurements of the polarization of the S phase can be used to predict ScS polarization in the epicentral distance interval between $60°$ and $70°$. Since the backazimuth is always zero for our source-receiver configuration (Figure 1a) and the radial ScS component is approximately a sign-flipped version of the S radial component (Figure 2), the sum of the S and the ScS polarizations must be approximately zero.

Figure 9 explores how accurately, under the assumptions described above, ScS splitting due to source-side anisotropy can be predicted from the splitting of the corresponding S phase. We do not incorporate any receiver-side or deep mantle anisotropy in these simulations. We measure S and ScS polarizations and splitting parameters (ϕ, δt). Then, we determine the difference between the ScS splitting parameters and those predicted from the S phase. Figure 9a shows an example for olivine anisotropy in the source-side upper mantle, with the elastic tensor rotated by $60°$ around the vertical axis (Section 2.1). The measured polariza-

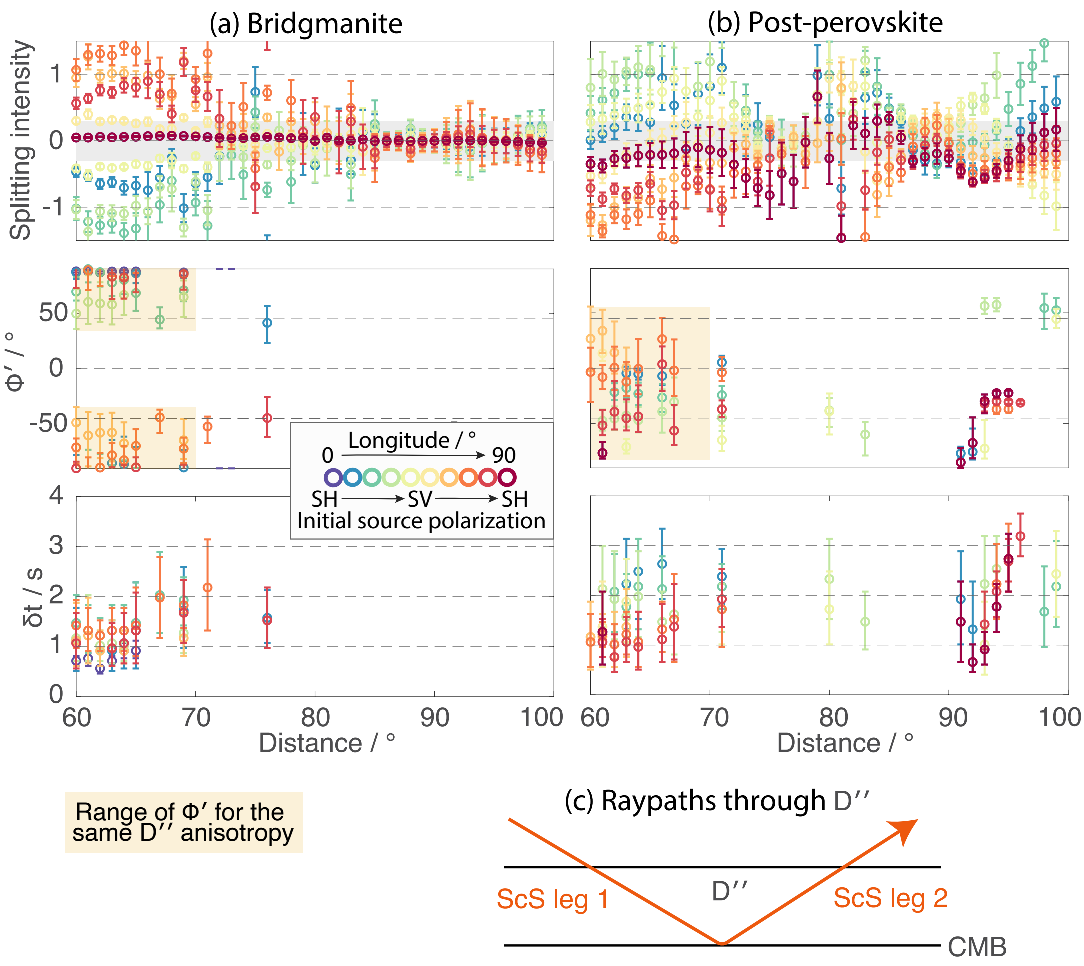


Figure 8 Shear wave splitting parameters for input models that include global layers of lowermost mantle anisotropy and no seismic anisotropy in the upper mantle. (a) Results for a $150\,\mathrm{km}$ thick layer of Br anisotropy at the base of the mantle and PREM-velocities otherwise. Same plotting conventions as in Figure 6a. The range of ϕ' values for the same deep mantle anisotropy, sampled from the same direction, but for different ScS initial source polarizations (see legend), is indicated by orange shading. (b) Results for a $175\,\mathrm{km}$ thick layer of Ppv anisotropy at the base of the mantle and PREM-velocities otherwise. Same plotting conventions as in panel (a). (c) Schematic illustration of the two ScS raypath legs through D′′. Because the measured shear wave splitting is a combination of the sampled seismic anisotropy on both legs of the raypath and the initial source polarization, a range of ϕ' values is obtained when sampling the same seismic anisotropy from the same direction for differently polarized ScS waves.

tions and fast polarization directions of ScS are different than those predicted from S by up to $\sim 35°$ and δt differs by up to $2\,\mathrm{s}$. Summary histograms for all the results obtained using a range of different rotation angles for the anisotropy geometry (Section 2.1) are shown in Figure 9b. These results indicate that substantial differences between predicted source-side anisotropy associated splitting parameters from S and measured splitting parameters for ScS are common. Also, the assumption of a radial component sign-flip of ScS caused by the reflection is imperfect, sometimes leading to polarization differences of up to $50°$.

Next, we systematically apply the source-side anisotropy splitting parameters, as inferred from S, to the ScS phase and then measure ScS splitting. If the source-side anisotropy correction was accurate, we would expect to measure null residual splitting from ScS as we did not incorporate deep mantle anisotropy in our simulations. (Recall that these simulations only include upper mantle anisotropy near the source.)

We define null measurements here as splitting measurements which have δt-values smaller than $0.3\,\mathrm{s}$, or a 95% confidence interval that overlaps with values $< 0.3\,\mathrm{s}$. This definition leads to few well-constrained (ϕ, δt)-measurements with $\delta t < 0.5\,\mathrm{s}$ in Figure 10. We find that for the HTI elastic tensor, only $\sim 63\,\%$ of the measured ScS splitting parameters are null after applying the source-side correction (Figure 10a). For olivine, this value is only $\sim 23\,\%$, meaning that in $\sim 77\,\%$ of the cases apparent D′′ splitting is introduced by applying the source-side correction (Figure 10b).

The reason that the source-side anisotropy correction is not generally accurate is that the source-side contribution for ScS cannot accurately be inferred from S. In Figure 11 we show retrieved δt values for well-constrained ScS splitting measurements after accounting for source-side anisotropy inferred from S. If ScS source-side splitting parameters are used to correct ScS, in $98\,\%$ of the cases no apparent D′′ splitting is introduced (Figure 11a), showing that our correction pro-

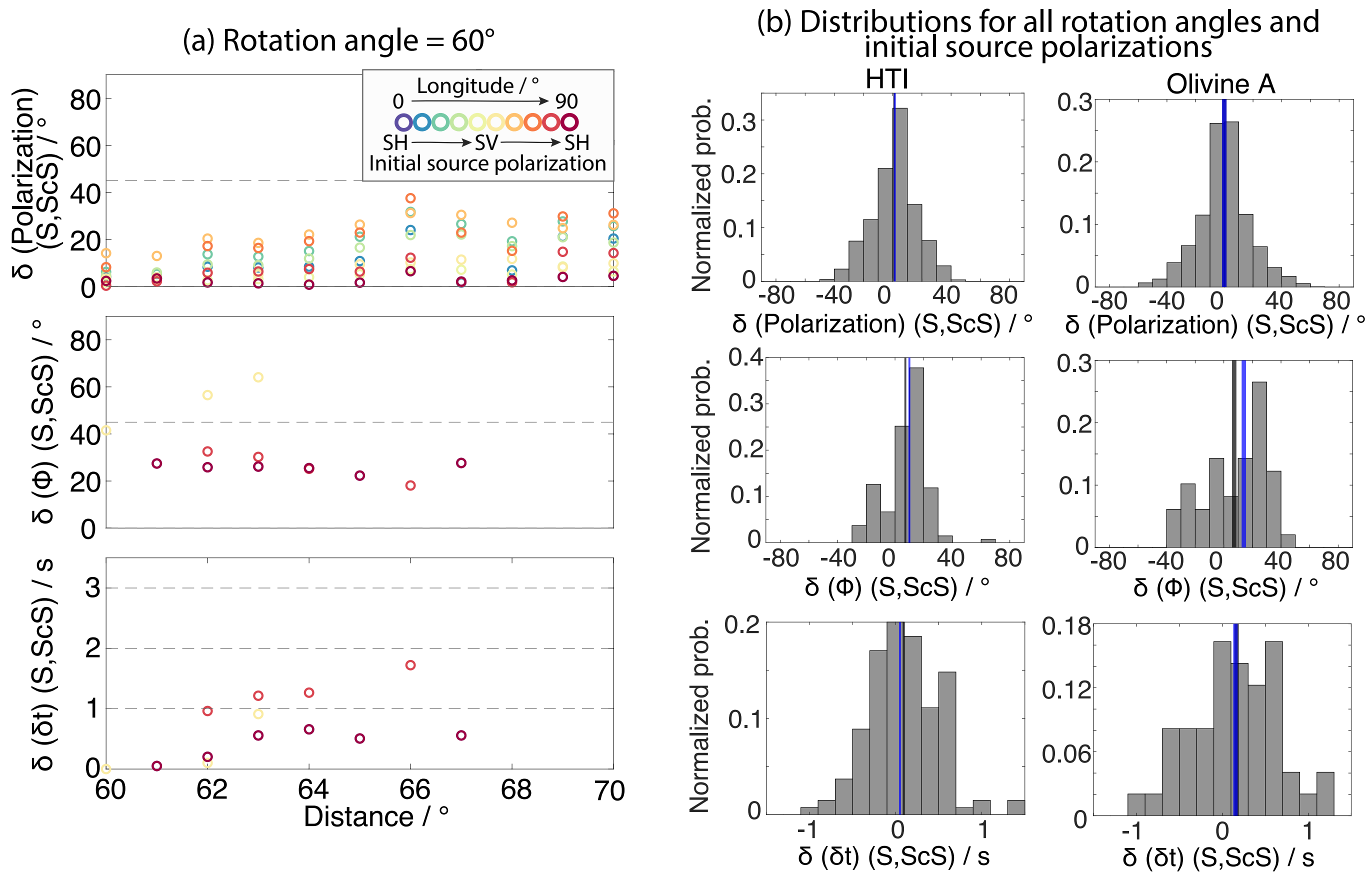


Figure 9 Differences between estimated ScS polarizations, measured fast polarization directions ϕ, and delay times δt, compared to the values that would be expected after S analysis if the radial ScS component was a perfectly sign-flipped version of the S radial component. Values are measured from synthetic waveforms generated for a model that includes source-side anisotropy only and PREM velocity structure elsewhere. Only well-constrained splitting results are shown (as defined in Section 5). (a) Example for the olivine elastic tensor shown in Figure 1c with a horizontal rotation angle of $60°$ (Section 2), and for different initial source polarizations (legend). Top: Absolute differences in predicted and measured ScS polarizations $\delta(Polarization)$. Middle: Absolute differences in predicted and measured ScS fast polarization directions $\delta(\phi)$. Bottom: Absolute differences in predicted and measured ScS delay times $\delta(\delta t)$. (b) Histograms showing the distributions of $\delta(Polarization)$ (top), $\delta(\phi)$ (middle) and $\delta(\delta t)$ (bottom) for HTI (left) and olivine (right) anisotropy (Figure 1c) for all initial initial polarizations and elastic tensor rotation angles (Section 2.1). Distribution means are shown as solid black lines and medians as solid blue lines.

cedure works well if splitting is perfectly known. We can use these insights to suggest three different strategies for accounting for source-side anisotropy. First, we can restrict measurements of ScS splitting to cases for which S source-side splitting is null (Figure 11b). Second, we can minimize the influence of source-side upper mantle anisotropy by only measuring ScS splitting from deep seismic events. However, the presence of seismic anisotropy has been suggested in the uppermost lower mantle, particularly in subduction zones has been suggested by several studies (e.g., Foley and Long, 2011; Lynner and Long, 2015; Mohiuddin et al., 2015). Therefore, such an approach would not necessarily (always) be reliable. Third, we can apply a source-side anisotropy correction if we measure a ScS polarization that is within $10°$ of the expected initial polarization for a sign-flip of the ScS radial component Figure 11c). In $90\,\%$ of these cases, null D$''$ splitting is correctly predicted from ScS if measured S source-side splitting is null (Figure 11b), suggesting that these strategies allow for the accurate consideration of source-side splitting in certain cases. In contrast, explicit source-side anisotropy corrections are inaccurate when these conditions are not met, even when ensuring that the ScS polarization is as expected from S (Figure 11c).

8 Correction of ScS for receiver-side anisotropy contribution inferred from SKS

As discussed in Section 7, we now better understand the strengths and weaknesses of the source-side correction for shear wave splitting; however, the receiver-side corrections remain to be explored. We next use a realistic synthetic setup to test the accuracy of receiver-side corrections. We again incorporate olivine A-type (Figure 1c) anisotropy in the upper mantle and infer upper mantle shear wave splitting parameters from the SKS seismic phase recorded at an epicentral distance of $100°$. We fit a $\sin(2\theta)$-curve to the SKS SI values as a function of backazimuth, as commonly done for real

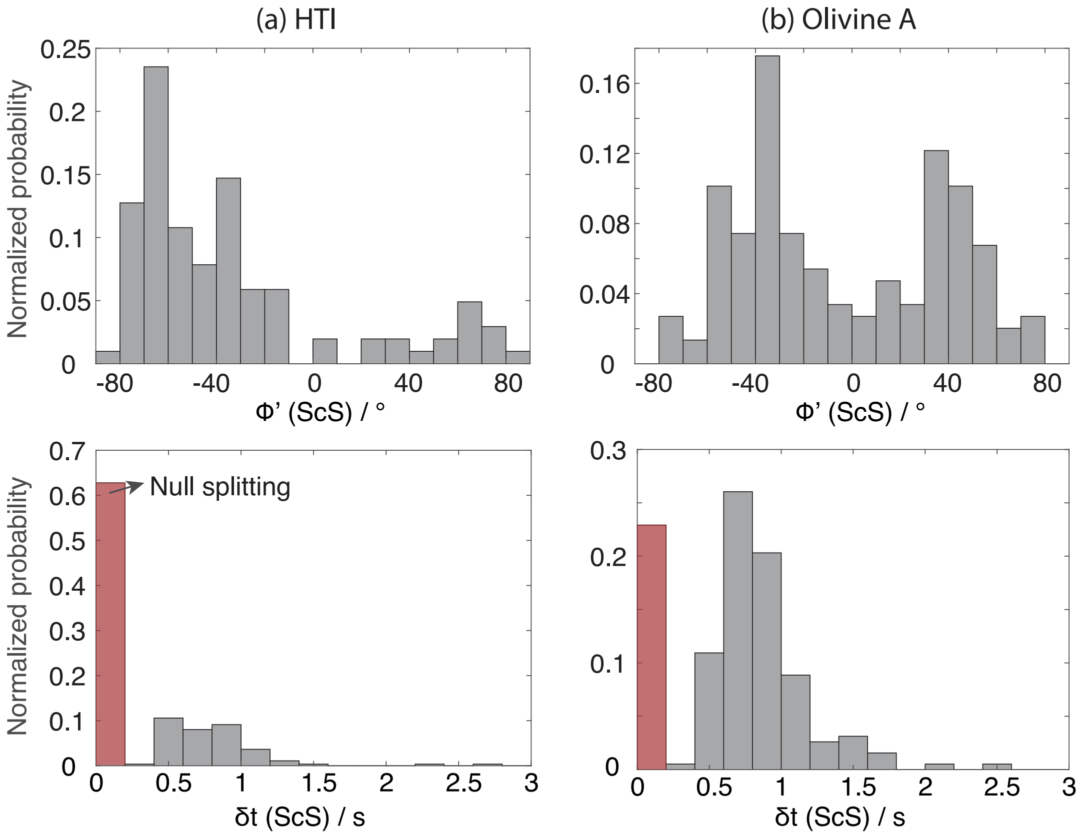


Figure 10 Distributions of measured ϕ' and δt for all initial polarizations and elastic tensor rotation angles after correcting for source-side anisotropy measured from S, only showing well-constrained results (as defined in Section 5). (a) Results for a model including HTI source-side anisotropy and PREM-velocities elsewhere. Top: ϕ' histograms; bottom: δt histograms. ϕ' values are only plotted for non-null measurements. Measurements are defined as null if the $95\,\%$ confidence interval of δt overlaps with the interval $[0\,\mathrm{s}, 0.3\,\mathrm{s}]$. Null measurements are indicated in orange color. (b) Same as panel (a) for olivine A source-side anisotropy.

data (e.g., Chevrot, 2000; Monteiller and Chevrot, 2010). These results are shown in Figure 12a. The determined best-fitting splitting parameters are then used to correct ScS for the effect of anisotropy beneath the receiver for simulations that only include upper mantle receiver-side anisotropy. For approximately $40\,\%$ of robust ScS measurements (for the setup shown in Figure 1) the measured splitting is null (Figure 12b), as expected. For the remaining $60\,\%$ of robust measurements, a variety of ϕ' and δt values are obtained (Figure 12b). This exercise demonstrates that explicit receiver-side corrections for upper mantle anisotropy beneath the receiver are likely unreliable in real data. The challenges are likely to be particularly given that splitting patterns as a function of backazimuth are often substantially more complicated than in this simple synthetic scenario.

An example of a robust (but artificial) ScS splitting measurement obtained after correcting for receiver-side anisotropy determined using SKS phases is shown in Figure 13. This particular case corresponds to a scenario in which the olivine A-type elastic tensor is sampled from a backazimuth of $80°$ in Figure 12a. It is apparent that the $\sin(2\theta)$-fit is imperfect for this backazimuth (Figure 12a), which is why the corrected waveform (Figure 13a) appears substantially split, the particle motion (Figure 13b) mimics splitting, and splitting parameters are well-constrained with very tight uncertainty intervals (Figure 13c), despite the lack of D″ anisotropy in this simulation.

While we conduct these measurements for ScS in this work, our calculations are similarly applicable for the measurement of S splitting after correcting for receiver-side anisotropy inferred from SKS, which is commonly done to infer seismic anisotropy in the transition zone in subduction zones (e.g., Russo et al., 2010; Mohiuddin et al., 2015; Eakin et al., 2018). One potential way to deal with this challenge may be to correct for upper mantle splitting beneath the receiver determined using other phases measured at the same backazimuth, preferably for the same source-receiver configuration. However, this appears challenging for ScS distance between $60°$ and $70°$, as there is no obvious additional phase that could be used for such an approach.

9 Discussion

9.1 How to infer D″ anisotropy from ScS splitting measurements

We have shown that D″ anisotropy is challenging to infer from ScS waves that arrive at epicentral distances $< 60°$, because CMB reflection coefficients for the SV component will strongly depend on the deep mantle velocity structure close to the ScS reflection point (Fig-

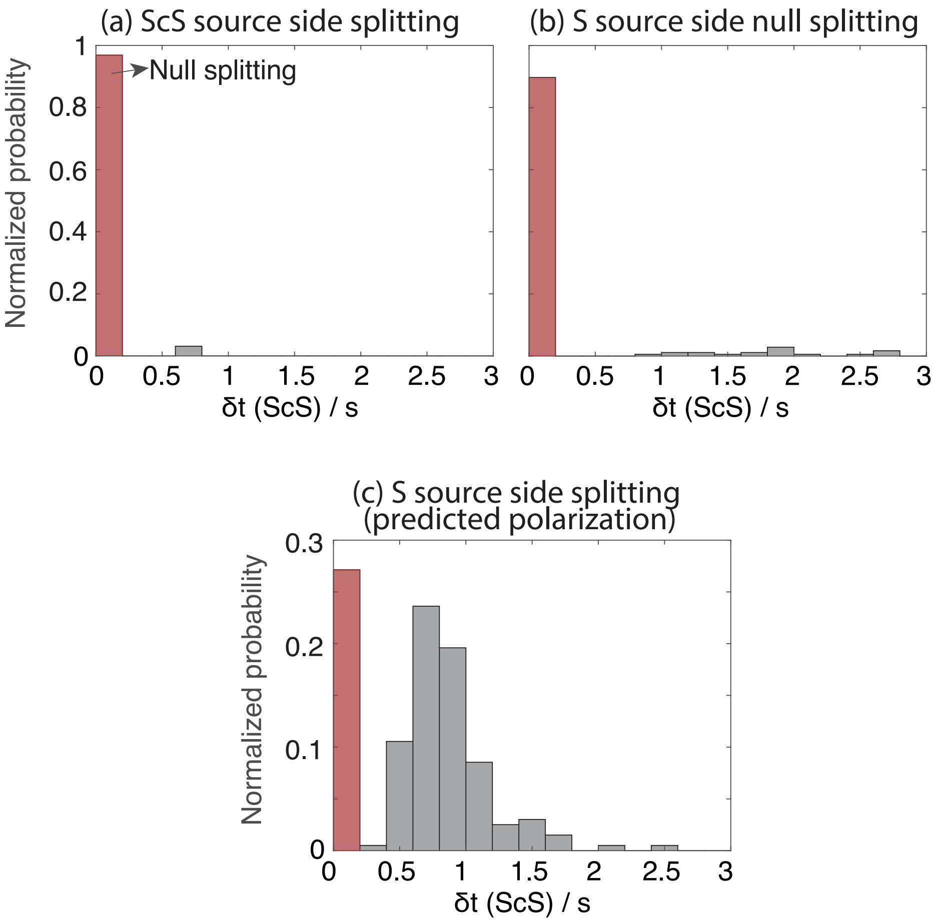


Figure 11 Retrieved δt values for well-constrained ScS splitting measurements after accounting for olivine source-side anisotropy, for the same set of simulations used in Figure 10. Only well-constrained splitting results are shown (as defined in Section 5). (a) Histograms of results obtained after correcting ScS waveforms for source-side anisotropy measured from the same ScS waveform. Plotting conventions are the same as in the bottom panel of Figure 10a. (b) ScS δt measurements for those S-ScS pairs for which S source-side splitting is null. Plotting conventions are as in panel (a). (c) Histogram of ScS δt measurements obtained after correcting for source-side anisotropy inferred from S for those S-ScS pairs for which ScS polarizations are as predicted from S (under the assumption of a sign-flip of the radial component of ScS through the CMB reflection). Plotting conventions are as in panel (a).

ure 3). Therefore, polarization directions of ScS, as well as apparent Pol/Pol_{90} amplitude ratios, will be influenced by effects other than seismic anisotropy. For epicentral distances between $70°$ and $80°$, the ScS arrival may potentially be contaminated by (P)PS or SP (Figures 2 and 5), which strongly influences ScS polarizations (Figure 5). This can, in some cases, cause apparent ScS splitting in absence of seismic anisotropy (Figure 6c). For even larger distances, S and ScS merge (Figures 2, 4 and 5), which can lead to effects that mimic splitting, even for simple isotropic models such as isotropic PREM (Figure 6). Apparent splitting caused by isotropic effects at these distances can be indistinguishable from shear wave splitting caused by lowermost mantle anisotropy, with the waveform shape of the Pol_{90} component approximately agreeing with the time derivative of the Pol_0 component (Figure 7). Therefore, we suggest that ScS shear wave splitting measurements are difficult to reliably perform for epicentral distances $> 70°$ and for most epicentral distances $< 60°$ (with the exception, perhaps, of almost vertical incidence angles for small distances).

In the candidate epicentral distance range between $60°$ and $70°$ for ScS splitting measurements, the receiver- and source-side anisotropy influence is often explicitly corrected to extract the lowermost mantle contribution. However, we have shown that explicit upper mantle anisotropy corrections can be unreliable (Figures 10 to 13) and therefore recommend only using ScS waves for which both source-side and receiver-side anisotropy are null. Practically, this means that ScS splitting measurements should only be applied at null stations for S-ScS pairs for which S source-side splitting is null. While these precautions mean that a much smaller number of S-ScS pairs are available for D″ splitting studies, they are likely to result in significantly higher-quality estimates of ScS splitting due to lowermost mantle anisotropy.

9.2 Global measurements of ScS splitting due to deep mantle anisotropy

We apply our strategy for estimating D″-associated ScS splitting measurements worldwide at the null stations reported by Lynner and Long (2013) and Walpole et al.

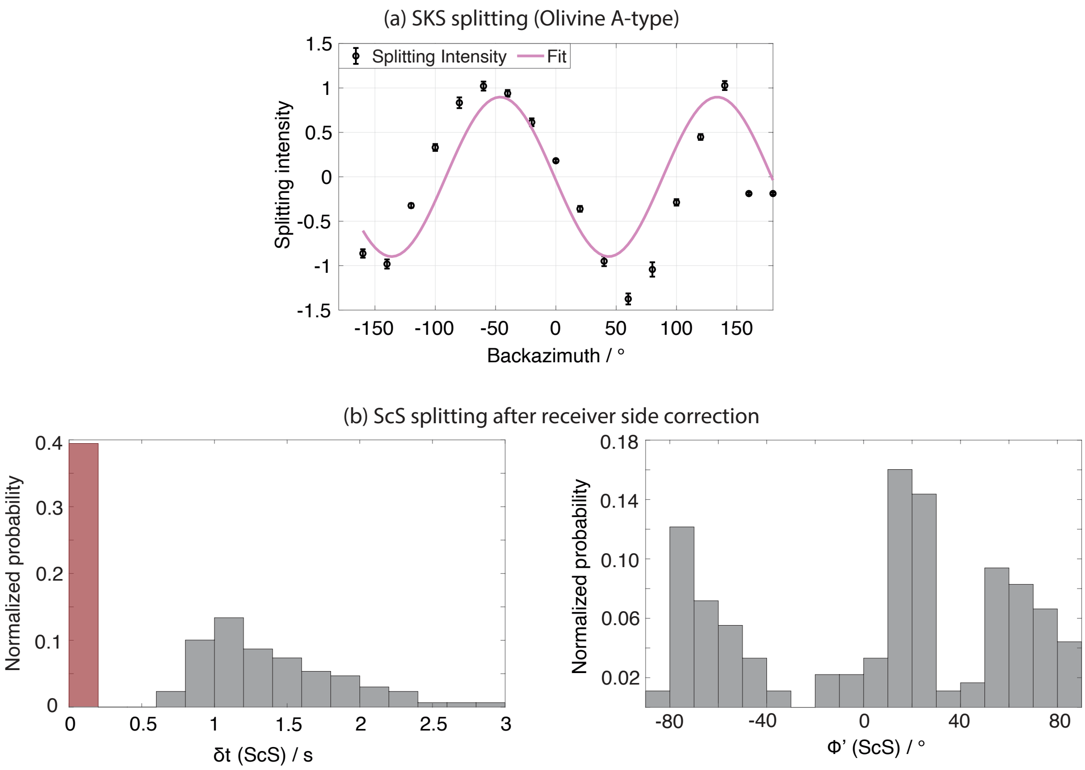


Figure 12 Apparent shear-wave splitting from ScS waveforms that were corrected for upper mantle anisotropy. (a) SKS SI values (black circles) and 95% confidence intervals (error bar) as a function of backazimuth. The fitted $\sin(2\theta)$-curve to determine best-fitting splitting parameters is shown in pink. The best-fitting values are ($\phi = -1.2°$, $\delta t = 0.9\,\mathrm{s}$). (b) ϕ' and δt histograms using the same plotting conventions as in Figure 10a.

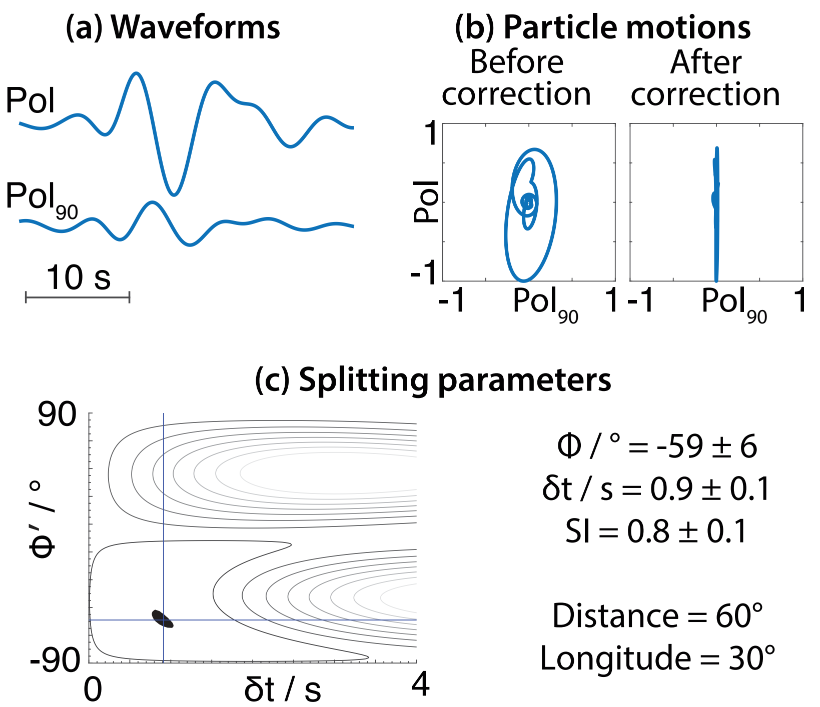


Figure 13 Measured shear wave splitting from waveforms corrected for upper mantle anisotropy inferred from SKS (Figure 12). Plotting conventions are identical to Figure 7.

(2014). We use all seismic events with moment magnitude > 5.7 in an epicentral distance range between $60°$ and $70°$ that occurred after January 1, 1990. The raypath coverage for all source-receiver pairs for which we could obtain well-constrained ScS splitting measurements is shown in Figure 14. Following the recommendations developed here, we only interpret ScS splitting as being indicative of deep mantle anisotropy if the S phase for the same source-receiver pair is not split. An example splitting measurement for such a S-ScS pair is presented in Figure 15.

We follow our suggested procedure to calculate split-

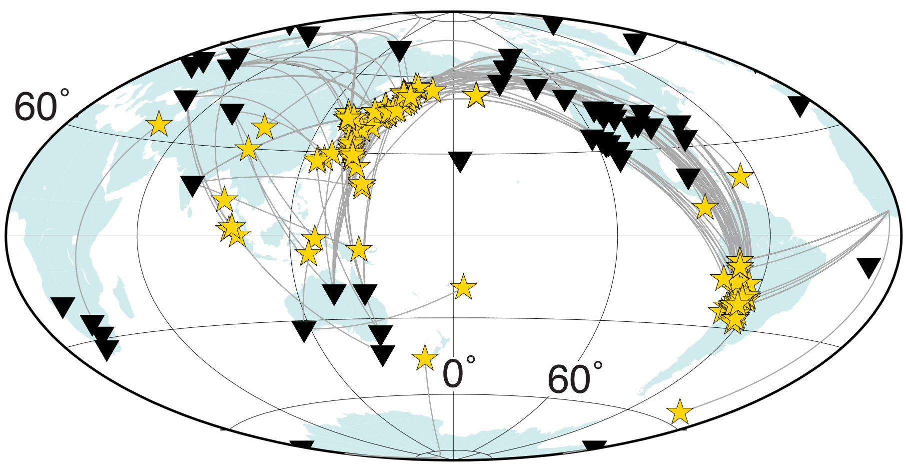


Figure 14 Source (yellow stars)-receiver (black triangles) configuration for global deep mantle anisotropy analysis using ScS. The selected stations are the null stations reported by Lynner and Long (2013) and Walpole et al. (2014). We only show sources for which well-constrained D″ anisotropy associated SI values could be obtained. Great circle raypaths are shown as gray lines.

ting parameters to all our seismic data for null stations. All ScS splitting measurements due to lowermost mantle anisotropy are shown in map view in Figure 16. In some cases only well-constrained splitting intensities can be obtained. In other cases, splitting parameters (ϕ', δt) can also be reliably measured. We can identify four different deep mantle regions A-D that show at least some evidence for anisotropy (Figure 16). Overall, we find evidence for seismic anisotropy in all regions in which ray coverage is good, suggesting that lowermost mantle anisotropy is likely widespread. These regions include central Asia (A), southeast Asia (B), northeast Russia/Alaska (C), and the Caribbean (D). North of region A, multiple studies have previously reported seismic anisotropy in D″ (e.g., Wookey and Kendall, 2008; Creasy et al., 2021). Our results for this raypath corridor approximately agree with the ϕ' values of $35°$ reported by Creasy et al. (2021) but are different than those from Wookey and Kendall (2008), who reported $\phi' \approx -7°$. However, as mentioned above, ϕ' values depend on the initial polarization of the ScS wave, which is why we cannot necessarily expect to obtain the same ϕ' values for a particular region if the ScS initial polarizations in the dataset vary. Grund and Ritter (2018), Thomas and Kendall (2002) and Wolf et al. (2022b) also identified lowermost mantle anisotropy in some parts of region A using a different methodology. These measurements are hard to directly compare with ours; however, these studies are consistent with our finding of D″ anisotropy here.

Deep mantle anisotropy in region B has not been previously studied. We find the lowermost mantle in this region to be generally anisotropic; however, the strength of splitting due to seismic anisotropy varies (Figure 16). ϕ'-values tend to be close to $0°$ in most cases, but this – on its own – is an insufficient constraint on the geometry of anisotropy without taking into account the wave's initial polarization. The ScS raypaths through D″, shown in the inset for region B (Figure 16), are close to the edge of the Pacific LLVP and show evidence for seismic anisotropy. This agrees with the finding of other studies that seismic anisotropy is often strong close to such edges (e.g., Wang and Wen, 2004; Lynner and Long; 2014; Deng et al., 2017, Reiss et al.; 2019; Wolf and Long; 2023).

Much of region C has been found to be anisotropic in previous studies (e.g., Wookey et al., 2005; Asplet et al., 2020, 2023; Suzuki et al., 2021; Wolf and Long; 2022; Wolf et al., 2023a; 2024). Direct comparisons to many of these studies are difficult because they used different methods to infer the presence of seismic anisotropy. Wookey et al. (2005) used S-ScS differential splitting to investigate the west portion of region C. Additionally, seismic anisotropy has been detected in the western part of region C using multiple different methods (e.g., Wolf and Long, 2022; Asplet et al., 2023), which include S-ScS and SKS-SKKS differential splitting as well as S_{diff} splitting. Notably, we also detect particularly strong seismic anisotropy in this region.

Seismic anisotropy in region D has been identified by a large number of previous studies (e.g., Kendall and Silver, 1996; Rokosky et al., 2004, 2006; Garnero et al., 2004; Maupin et al., 2005; Nowacki et al., 2010). The study by Nowacki et al. (2010) also used S-ScS differential splitting measurements. Interestingly, we find splitting due to seismic anisotropy to be strong beneath central America and almost absent further to the east (Figure 16). In the northwest part of region D, shear wave splitting is weak as well, while it is substantially stronger in the southwest (Figure 16). To the east of region D we obtain five measurements that consistently show no evidence for splitting due to deep mantle anisotropy , and whose initial polarizations differ by up to $40°$. Therefore, we find the deep mantle in this region to likely be isotropic, in disagreement with the findings of Pisconti et al. (2023).

Due to the constraints that we impose in our approach to the measurement of ScS splitting, a large majority of seismograms cannot be used to reliably measure ScS splitting due to lowermost mantle anisotropy. A back-of-the-envelope calculation suggests that approximately 15 million three-component seismograms are currently

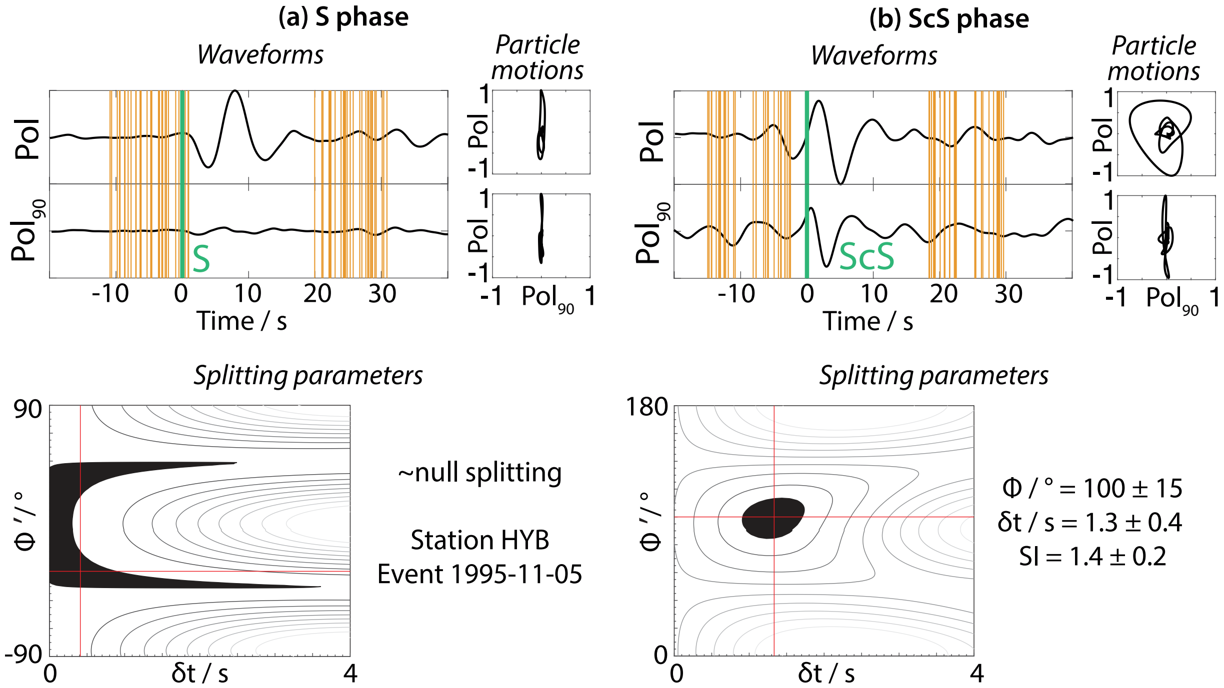


Figure 15 Example of an S and ScS splitting pair used to infer deep mantle anisotropy at station HYB for an event of 1995-11-05. (a) S Pol and Pol_{90} component velocity waveforms (top left). The PREM-predicted phase arrival time is shown by a light green line and the start/end of the automatically selected measurement windows are shown as orange lines. At the right we show S particle motions before (top) and after (bottom) correcting for the best fitting splitting parameters. At bottom, we show the best fitting splitting parameters for S in the ϕ'-δt plane. The 95% confidence interval is shown in black. Splitting for this S arrival is null. (b) Same as panel (a) for the corresponding ScS phase. ScS splitting parameters are shown in the bottom right.

publicly available for seismic events with moment magnitudes over 6.0. In this work, we obtain $\sim$130 robust ScS splitting measurements for seismic events with such moment magnitudes, using all null stations known to us (which may not be all that exist). Following this line of reasoning, under the constraints used in this study, only one out of every 100,000 seismograms is expected to yield a robust ScS splitting measurement of lowermost mantle anisotropy – a very small minority of available data. However, with the suggestions we put forward in Section 10, it may be possible increase this number.

9.3 Interpretations of ScS splitting measurements due to deep mantle anisotropy

Our work demonstrates that when multiple sets of splitting parameters (ϕ, δt) due to lowermost mantle anisotropy can be estimated in a particular region, a significant spread of these values can be expected (Figure 8). The reason is that the measured (ϕ, δt)-values do not only depend on the nature of deep mantle anisotropy but also on the initial polarization of ScS. Therefore, the measurement scatter shown in Figure 16 does not imply that measurements are unreliable because displaying measurements in map projection does not account for the wave's initial polarization. In fact, all SI measurements that are plotted on top of each other in Figure 16 and whose SI values differ have at least somewhat different initial polarizations. In order to thoroughly characterize the geometry of deep mantle seismic anisotropy measured from ScS waves, a sufficient number of splitting measurements is needed to allow for forward modeling or inversions that explicitly consider the initial polarization of each wave. Practically, a sufficient number of measurements may be difficult to obtain in many regions, given the substantial restrictions imposed to correctly account for potential upper mantle contributions.

One potential way to make use of ScS splitting measurements to constrain the geometry of anisotropy as opposed to merely using them as an anisotropy detector is to interpret them along with other independent constraints, such as SKS-SKKS differential splitting (e.g., Asplet et al., 2023), D″ reflection polarities (e.g., Pisconti et al., 2023), or S_{diff} splitting (e.g., Wolf et al., 2023b). When multiple constraints are available, ray-theoretical forward modeling (e.g., Wolf et al., 2019; Pisconti et al., 2023), full-wave simulations (e.g., Wolf et al., 2022a; 2022b), or inversions of ScS waveforms (Asplet et al., 2023) can potentially shed light on deformation in the deep mantle.

10 Ways forward

We have shown that it can be difficult to infer deep mantle anisotropy from ScS splitting measurements due to potential contamination from upper mantle anisotropy, which is difficult to account for. We have suggested a strategy of avoiding explicit upper mantle anisotropy corrections going forward by focusing on null stations and S-ScS pairs for which S is not split due to source-side

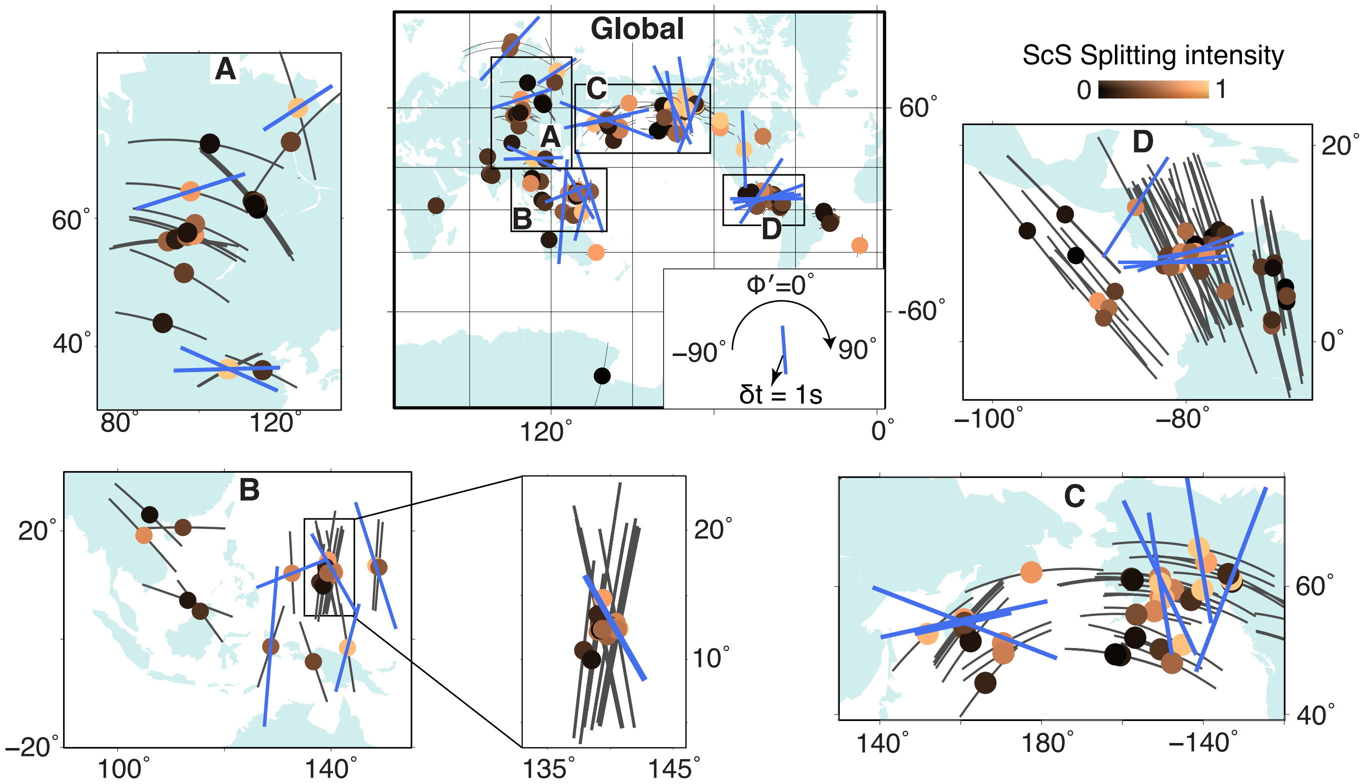


Figure 16 ScS splitting results for S-ScS pairs measured for station-event pairs shown in Figure 15, for which S splitting is null. Top center: Global results. ScS great circle raypaths through the lowermost 250 km of the mantle are shown as dark gray lines. Well-constrained ScS splitting intensity values are shown as colored circles (legend) at the ScS reflection point at the CMB. Splitting parameters (ϕ', δt; see legend) are shown as blue sticks and centered at the ScS reflection location. Different regions A-D with many well-constrained measurements are identified and labeled. Zoom-ins to these regions are shown in panels A-D surrounding the central plot.

upper mantle anisotropy. Crucial for this approach will be the identification of more null stations worldwide. In this work, we have used the null stations identified by Lynner and Long (2013) and Walpole et al. (2014); however, more null stations likely exist. Given the increased availability of seismic data since these two studies were published, it appears worthwhile to automatically and uniformly analyze all available broadband data to search for null stations, for example using an approach similar to Walpole et al. (2014).

Another possibility to increase the number of ScS splitting measurements due to deep mantle anisotropy is to use beamforming, which has only recently been applied in shear wave splitting studies (Wolf et al., 2023a). It has been shown that a beamforming approach effectively averages the upper mantle anisotropy contribution across the individual stations used to construct the beam (Wolf et al., 2023a). Therefore, it is possible to intentionally select stations such that the upper mantle anisotropy contribution to the beam beneath the array on the receiver side is effectively null. For such a station configuration, ScS splitting can be measured if the corresponding S beam splitting for the same source-array combination is null, indicating the absence of source-side anisotropy.

Interpretations of ScS splitting results in terms of anisotropic geometry will continue to be challenging. For such interpretations, the initial polarization of ScS will have to be explicitly considered. This has effectively been done by Asplet et al. (2023) by ensuring that ScS polarizations (approximately) agree with the backazimuth (through analysis of particle motions), although they used explicit upper mantle anisotropy corrections in their approach. At least in theory one could even go further: Seismic anisotropy in the lowermost mantle could be characterized by analyzing splitting intensities as a function of initial polarization for waves that sample the same lowermost mantle portion. However, given that we are dealing with two-layer splitting, this requires a much larger number of measurements than we have obtained for any particular region in this study (Figure 16). Most previous studies have not explicitly taken into account the ScS polarization and operated under the assumption that splitting due to D′′ anisotropy should lead to the same (ϕ', δt) values for the same region and sampling direction (e.g., Creasy et al., 2017; Wolf et al., 2019; Pisconti et al., 2023).

If a sufficient number of ScS splitting measurements from earthquakes with different initial source polarizations can be obtained for a given set of raypaths sampling D′′, a two-layer inversion for splitting parameters on the down- and upgoing leg of the raypath appears promising. Such an approach can be applied analogously to two-layer splitting analysis for the upper mantle (e.g., Silver and Savage, 1994; Wolfe and Silver, 1998; Aragon et al., 2017; Link et al., 2022). In our study, unfortunately, the number of well-constrained (ϕ, δt) measurements in any particular region is insufficient for the implementation of such an approach. As mentioned above, SI scattering is often straightforward to explain by different initial polarizations; in contrast, a precise characterization of the seismic anisotropy is challenging unless a large number of SI values for the same region can be obtained. Much easier is the detection of isotropic regions through initial polarization analysis, such as the isotropic region east of region D. The reason is that no more than a handful of null measurements with mutually different initial polarizations need to be obtained for the reliable characterization of an isotropic lowermost mantle region.

Going forward, it will also be important to combine ScS constraints with constraints from other seismic phases, whether waveform inversions (e.g., Asplet et al., 2023) or ray-theoretical forward modeling (e.g., Ford et al., 2015; Wolf et al., 2019; Pisconti et al., 2023) approaches are used. Given the issues that have been pointed out with the use of ray-theoretical assumptions (Nowacki and Wookey, 2016; Wolf et al., 2022a), it will be preferable to move away from ray-theory in future studies and make use of available full-wave modeling tools, including AxiSEM3D.

11 Conclusions

Using global wavefield simulations and calculations of ScS reflection coefficients, we have explored how ScS polarizations are affected by the CMB reflection. We find that measured ScS polarizations at the receiver, depend not only on the initial source polarization, but also on the deep mantle velocity structure at the reflection point and on the epicentral distance under consideration. In particular, in the epicentral distance between $60°$ and $70°$, the CMB reflection can be well approximated as a sign-flip of SV, while SH is unaltered. For distances close than $60°$, SV amplitudes are affected by the reflection, and for distances above $70°$, apparent shear wave splitting can be introduced for isotropic input models due to phase interference, for example with S. Therefore, the distance range suitable for ScS splitting measurements is $60°$ to $70°$.

If ScS shear wave splitting is caused by lowermost mantle anisotropy, the measured apparent splitting parameters are substantially influenced by the initial source polarization of the wave. The reason is that each leg of the ScS raypath through D″ (downgoing and upgoing) experience splitting separately. Therefore, for any D″ region that is sampled from the same direction by multiple ScS waves, we would expect to measure a range of apparent splitting parameters that depend on the initial polarizations of the ScS waves. We have shown that if an anisotropy contribution on the source side is inferred from S splitting and then used to correct the ScS waveform, in many cases apparent D″ splitting can be introduced. Similar issues exist for explicit receiver-side corrections. Therefore, we suggest a strategy that only uses null stations to infer deep mantle anisotropy from ScS. Measurements of ScS splitting at null stations should only be attributed to deep mantle anisotropy if the measured S splitting for the same source-receiver pair is null. We have applied this analysis strategy globally and detected deep mantle seismic anisotropy in multiple regions around the Earth, including regions that have not been shown to be anisotropic before, for example, southern Russia and the southwestern Pacific Ocean.

Going forward, to improve D″ anisotropy sampling using ScS, the identification of more null stations and the implementation of beamforming approaches in terms of the geometry of anisotropy will be helpful. Interpretations of ScS splitting going beyond using ScS as a simple anisotropy detector will need to consider the initial polarization of each ScS wave as well as potentially different splitting on the two ScS raypath legs through D″. While this approach is not typically incorporated in ScS splitting studies at present, it holds promise for gaining insight into the geometry of the anisotropy, and thus flow at the base of the mantle.

Acknowledgements

This work was funded by Yale University and by the U.S. National Science Foundation via grant EAR-2026917 to MDL. We thank the Yale Center for Research Computing for providing the necessary research computing infrastructure for this study. This research was inspired by discussions with Ed Garnero. We thank Sanne Cottaar and the editor, Lauren Waszek, for constructive reviews.

Data and code availability

All data used in this study are publicly available through Earthscope (http://service.iris.edu), SCEDC (http://service.-scedc.caltech.edu), RESIF (http://ws.resif.fr) and GEOFON (http://geofon.gfz-potsdam.de) data centers. We used data from USArray (IRIS Transportable Array, 2003) and data from networks AK (Alaska Earthquake Center, Univ. of Alaska Fairbanks, 1987), AT (NOAA National Oceanic and Atmospheric Administration (USA), 1967), CI (California Institute of Technology and United States Geological Survey Pasadena, 1926), CN (Canada), 1975), G (Institut de physique du globe de Paris (IPGP) and École et Observatoire des Sciences de la Terre de Strasbourg (EOST), 1982), GE (GEOFON Data Centre, 1993), GT (Albuquerque Seismological Laboratory (ASL)/USGS, 1993), II (Scripps Institution of Oceanography, 1986), IU (Albuquerque Seismological Laboratory/USGS, 2014), NL (KNMI, 1993), NR (Utrecht University (UU Netherlands), 1983), PM (Instituto Português do Mar e da Atmosfera, I.P., 2006), IU (Albuquerque Seismological Laboratory/USGS, 2014). The synthetic seismograms for this study were computed using AxiSEM3D, which is publicly available at https://github.com/AxiSEMunity (Fernando et al., 2024). The Generic Mapping Tools (Wessel and Smith, 1998), ObsPy (Beyreuther et al., 2010), MSAT (Walker and Wookey, 2012), SplitRacer (Reiss and Rümpker, 2017) and AxiSEM3D (Leng et al., 2016) were used in this research.

Competing interests

The authors declare no competing interests with respect to this manuscript.

References

Alaska Earthquake Center, Univ. of Alaska Fairbanks. Alaska Geophysical Network, 1987. doi: 10.7914/SN/AK.

Albuquerque Seismological Laboratory (ASL)/USGS. Global Telemetered Seismograph Network (USAF/USGS), 1993. doi: 10.7914/SN/GT.

Albuquerque Seismological Laboratory/USGS. Global Seismograph Network (GSN - IRIS/USGS), 2014. doi: 10.7914/SN/IU.

Aragon, J. C., Long, M. D., and Benoit, M. H. Lateral Variations in SKS Splitting Across the MAGIC Array, Central Appalachians. *Geochemistry, Geophysics, Geosystems*, 18:4136–4155, 2017.

Asplet, J., Wookey, J., and Kendall, M. A potential post-perovskite province in D′′ beneath the Eastern Pacific: evidence from new analysis of discrepant SKS–SKKS shear-wave splitting. *Geophysical Journal International*, 221:2075–2090, 2020. doi: 10.1093/gji/ggaa114.

Asplet, J., Wookey, J., and Kendall, M. Inversion of shear wave waveforms reveal deformation in the lowermost mantle. *Geophysical Journal International*, 232:97–114, 2023. doi: 10.1093/gji/ggac328.

Barruol, G. and Kern, H. Seismic anisotropy and shear-wave splitting in lower-crustal and upper-mantle rocks from the Ivrea Zone: experimental and calculated data. *Physics of the Earth and Planetary Interiors*, 95:175–194, 1996. doi: 10.1016/0031-9201(95)03124-3.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., and Wassermann, J. ObsPy: A Python Toolbox for Seismology. *Seismological Research Letters*, 81:530–533, 06 2010. doi: 10.1111/10.1785/gssrl.81.3.530.

Borgeaud, A. F., Konishi, K., Kawai, K., and Geller, R. J. Finite frequency effects on apparent S-wave splitting in the D′′ layer: comparison between ray theory and full-wave synthetics. *Geophysical Journal International*, 207:12–28, 2016. doi: 10.1093/gji/ggw254.

California Institute of Technology and United States Geological Survey Pasadena. Southern California Seismic Network, 1926. doi: 10.7914/SN/CI.

Canada), N. R. C. N. Canadian National Seismograph Network, 1975. doi: 10.7914/SN/CN.

Chang, S.-J. and Ferreira, A. M. G. Inference on Water Content in the Mantle Transition Zone Near Subducted Slabs From Anisotropy Tomography. *Geochemistry, Geophysics, Geosystems*, 20:1189–1201, 2019. doi: 10.1029/2018GC008090.

Chang, S.-J., Ferreira, A. M. G., Ritsema, J., van Heijst, H. J., and Woodhouse, J. H. Joint inversion for global isotropic and radially anisotropic mantle structure including crustal thickness perturbations. *Journal of Geophysical Research: Solid Earth*, 120 (6):4278–4300, 2015. doi: 10.1002/2014JB011824.

Chapman, C. *Fundamentals of Seismic Wave Propagation*. Cambridge University Press, 2004. doi: 10.1017/CBO9780511616877.

Chevrot, S. Multichannel analysis of shear wave splitting. *Journal of Geophysical Research: Solid Earth*, 105:21579–21590, 2000. doi: 10.1029/2000JB900199.

Cottaar, S. and Romanowicz, B. Observations of changing anisotropy across the southern margin of the African LLSVP. *Geophysical Journal International*, 195:1184–1195, 2013. doi: 10.1093/gji/ggt285.

Creasy, N., Long, M. D., and Ford, H. A. Deformation in the lowermost mantle beneath Australia from observations and models of seismic anisotropy. *Journal of Geophysical Research: Solid Earth*, 122:5243–5267, 2017. doi: 10.1002/2016JB013901.

Creasy, N., Miyagi, L., and Long, M. D. A Library of Elastic Tensors for Lowermost Mantle Seismic Anisotropy Studies and Comparison With Seismic Observations. *Geochemistry, Geophysics, Geosystems*, 21:e2019GC008883, 2020. doi: 10.1029/2019GC008883.

Creasy, N., Pisconti, A., Long, M. D., and Thomas, C. Modeling of Seismic Anisotropy Observations Reveals Plausible Lowermost Mantle Flow Directions Beneath Siberia. *Geochemistry, Geophysics, Geosystems*, 22(10):e2021GC009924, 2021. doi: 10.1029/2021GC009924.

Deng, J., Long, M. D., Creasy, N., Wagner, L., Beck, S., Zandt, G., Tavera, H., and Minaya, E. Lowermost mantle anisotropy near the eastern edge of the Pacific LLSVP: constraints from SKS-SKKS splitting intensity measurements. *Geophysical Journal International*, 210:774–786, 2017. doi: 10.1093/gji/ggx190.

Dziewonski, A. M. and Anderson, D. L. Preliminary reference Earth model. *Physics of the Earth and Planetary Interiors*, 25:297–356, 1981. doi: 10.1016/0031-9201(81)90046-7.

Eakin, C., Rychert, C., and Harmon, N. The Role of Oceanic Transform Faults in Seafloor Spreading: A Global Perspective From Seismic Anisotropy. *Journal of Geophysical Research: Solid Earth*, 123:1736–1751, 2018.

Fernando, B., Wolf, J., Leng, K., Nissen-Meyer, T., Eaton, W., Styczinski, M., Walker, A., Craig, T., Muir, J., Nunn, C., and Long, M. AxiSEM3D - an introduction to using the code and its applications. *EarthArXiv*, 2024. doi: 10.31223/X5TH7P.

Foley, B. J. and Long, M. D. Upper and mid–mantle anisotropy beneath the Tonga slab. *Geophysical Research Letters*, 38, 2011. doi: 10.1029/2010GL046021.

Ford, H. A., Long, M. D., He, X., and Lynner, C. Lowermost mantle flow at the eastern edge of the African Large Low Shear Velocity Province. *Earth and Planetary Science Letters*, 420:12–22, 2015. doi: 10.1016/j.epsl.2015.03.029.

Fouch, M. J., Fischer, K. M., and Wysession, M. E. Lowermost mantle anisotropy beneath the Pacific: Imaging the source of the Hawaiian plume. *Earth and Planetary Science Letters*, 190(3): 167–180, 2001. doi: 10.1016/S0012-821X(01)00380-6.

French, S. W. and Romanowicz, B. A. Whole-mantle radially anisotropic shear velocity structure from spectral-element waveform tomography. *Geophysical Journal International*, 199 (3):1303–1327, 2014. doi: 10.1093/gji/ggu334.

Garnero, E. J. and Lay, T. Lateral variations in lowermost mantle shear wave anisotropy beneath the north Pacific and Alaska. *Journal of Geophysical Research: Solid Earth*, 102(B4): 8121–8135, 1997. doi: 10.1029/96JB03830.

Garnero, E. J., Maupin, V., Lay, T., and Fouch, M. J. Variable Azimuthal Anisotropy in Earth's Lowermost Mantle. *Science*, 306 (5694):259–261, 2004. doi: 10.1126/science.1103411.

GEOFON Data Centre. GEOFON Seismic Network, 1993. doi: 10.14470/TR560404.

Grund, M. and Ritter, J. R. Widespread seismic anisotropy in Earth's lowermost mantle beneath the Atlantic and Siberia. *Geology*, 47:123–126, 2018. doi: 10.1130/G45514.1.

Haws, A. A., Long, M. D., and Luo, Y. Anisotropic structure of the normally-dipping and flat slab segments of the Alaska subduction zone: Insights from receiver function analysis. *Tectonophysics*, 868:230112, 2023. doi: 10.1016/j.tecto.2023.230112.

Institut de physique du globe de Paris (IPGP) and École et Observatoire des Sciences de la Terre de Strasbourg (EOST). GEOSCOPE, French Global Network of broad band seismic stations, 1982. doi: 10.18715/GEOSCOPE.G.

Instituto Português do Mar e da Atmosfera, I.P. Portuguese National Seismic Network, 2006. doi: 10.7914/SN/PM.

IRIS Transportable Array. USArray Transportable Array, 2003. doi: 10.7914/SN/TA.

Karato, S.-i. *Deformation of Earth Materials: An Introduction to the Rheology of Solid Earth*. Cambridge University Press, 2008. doi: 10.1017/CBO9780511804892.

Kendall, J.-M. and Silver, P. Constraints from seismic anisotropy on the nature of the lower mantle. *Nature*, 381:409–412, 1996. doi: 10.1038/381409a0.

KNMI. Netherlands Seismic and Acoustic Network, 1993. doi: 10.21944/E970FD34-23B9-3411-B366-E4F72877D2C5.

Komatitsch, D., Vinnik, L. P., and Chevrot, S. SHdiff-SVdiff splitting in an isotropic Earth. *Journal of Geophysical Research: Solid Earth*, 115(B7), 2010. doi: 10.1029/2009JB006795.

Leng, K., Nissen-Meyer, T., and van Driel, M. Efficient global wave propagation adapted to 3-D structural complexity: a pseudospectral/spectral-element approach. *Geophysical Journal International*, 207(3):1700–1721, 2016. doi: 10.1093/gji/ggw363.

Leng, K., Nissen-Meyer, T., van Driel, M., Hosseini, K., and Al-Attar, D. AxiSEM3D: broad-band seismic wavefields in 3-D global earth models with undulating discontinuities. *Geophysical Journal International*, 217(3):2125–2146, 2019. doi: 10.1093/gji/ggz092.

Lin, Y.-P., Zhao, L., and Hung, S.-H. Full-wave effects on shear wave splitting. *Geophysical Research Letters*, 41:799–804, 2014. doi: 10.1002/2013GL058742.

Link, F., Reiss, M. C., and Rümpker, G. An automatized XKS-splitting procedure for large data sets: Extension package for SplitRacer and application to the USArray. *Computers & Geosciences*, 158: 104961, 2022. doi: 10.1016/j.cageo.2021.104961.

Long, M. D. Complex anisotropy in D′′ beneath the eastern Pacific from SKS–SKKS splitting discrepancies. *Earth and Planetary Science Letters*, 283:181–189, 2009. doi: 10.1016/j.epsl.2009.04.019.

Long, M. D. and Becker, T. Mantle dynamics and seismic anisotropy. *Earth and Planetary Science Letters*, 297:341–354, 2010. doi: 10.1016/j.epsl.2010.06.036.

Lynner, C. and Long, M. D. Sub-slab seismic anisotropy and mantle flow beneath the Caribbean and Scotia subduction zones: Effects of slab morphology and kinematics. *Earth and Planetary Science Letters*, 361:367–378, 2013. doi: 10.1016/j.epsl.2012.11.007.

Lynner, C. and Long, M. D. Lowermost mantle anisotropy and deformation along the boundary of the African LLSVP. *Geophysical Research Letters*, pages 3447–3454, 2014. doi: 10.1002/2014GL059875.

Lynner, C. and Long, M. D. Heterogeneous seismic anisotropy in the transition zone and uppermost lower mantle: evidence from South America, Izu-Bonin and Japan. *Geophysical Journal International*, 201:1545–1552, 2015. doi: 10.1093/gji/ggv099.

Maupin, V., Garnero, E. J., Lay, T., and Fouch, M. J. Azimuthal anisotropy in the D′′ layer beneath the Caribbean. *Journal of Geophysical Research: Solid Earth*, 110(B8):B08301, 2005. doi: 10.1029/2004JB003506.

Mohiuddin, A., D. Long, M., and Lynner, C. Mid-mantle seismic anisotropy beneath Southwestern Pacific subduction systems and implications for mid-mantle deformation. *Physics of the Earth and Planetary Interiors*, 245, 2015. doi: 10.1016/j.pepi.2015.05.003.

Monteiller, V. and Chevrot, S. How to make robust splitting measurements for single-station analysis and three-dimensional imaging of seismic anisotropy. *Geophysical Journal International*, 182:311–328, 2010. doi: 10.1111/j.1365-246X.2010.04608.x.

Nissen-Meyer, T., van Driel, M., Stähler, S. C., Hosseini, K., Hempel, S., Auer, L., Colombi, A., and Fournier, A. AxiSEM: broadband 3-D seismic wavefields in axisymmetric media. *Solid Earth*, 5: 425–445, 2014. doi: 10.5194/se-5-425-2014.

Niu, F. and Perez, A. M. Seismic anisotropy in the lower mantle: A comparison of waveform splitting of SKS and SKKS. *Geophysical Research Letters*, 31, 2004. doi: 10.1029/2004GL021196.

NOAA National Oceanic and Atmospheric Administration (USA). National Tsunami Warning Center Alaska Seismic Network, 1967. doi: 10.7914/SN/AT.

Nowacki, A. and Cottaar, S. *Toward Imaging Flow at the Base of the Mantle with Seismic, Mineral Physics, and Geodynamic Constraints*, chapter 13, pages 329–352. American Geophysical Union (AGU), 2021. doi: 10.1002/9781119528609.ch13.

Nowacki, A. and Wookey, J. The limits of ray theory when measuring shear wave splitting in the lowermost mantle with ScS waves. *Geophysical Journal International*, 207:1573–1583, 2016. doi: 10.1093/gji/ggw358.

Nowacki, A., Wookey, J., and Kendall, J.-M. Deformation of the lowermost mantle from seismic anisotropy. *Nature*, 467: 1091–1094, 2010. doi: 10.1038/nature09507.

Panning, M. and Romanowicz, B. A three-dimensional radially anisotropic model of shear velocity in the whole mantle. *Geophysical Journal International*, 167:361–379, 2006. doi: 10.1111/j.1365-246X.2006.03100.x.

Parisi, L., Ferreira, A. M. G., and Ritsema, J. Apparent Splitting of S Waves Propagating Through an Isotropic Lowermost Mantle. *Journal of Geophysical Research: Solid Earth*, 123:3909–3922, 2018. doi: 10.1002/2017JB014394.

Pisconti, A., Creasy, N., Wookey, J., Long, M. D., and Thomas, C. Mineralogy, fabric and deformation domains in D′′ across the southwestern border of the African LLSVP. *Geophysical Journal International*, 232:705–724, 2023. doi: 10.1093/gji/ggac359.

Pulliam, J. and Sen, M. K. Seismic anisotropy in the core—mantle transition zone. *Geophysical Journal International*, 135: 113–128, 1998. doi: 10.1046/j.1365-246X.1998.00612.x.

Reiss, M. and Rümpker, G. SplitRacer: MATLAB Code and GUI for Semiautomated Analysis and Interpretation of Teleseismic Shear-Wave Splitting. *Seismological Research Letters*, 88:392 –– 409, 2017. doi: 10.1785/0220160191.

Reiss, M. C., Long, M. D., and Creasy, N. Lowermost Mantle Anisotropy Beneath Africa From Differential SKS-SKKS Shear-Wave Splitting. *Journal of Geophysical Research: Solid Earth*, 124(8):8540–8564, 2019. doi: 10.1029/2018JB017160.

Ritsema, J., Deuss, A., van Heijst, H. J., and Woodhouse, J. H. S40RTS: a degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements. *Geophysical Journal International*, 184(3):1223–1236, 2011. doi: 10.1111/j.1365-246X.2010.04884.x.

Rokosky, J. M., Lay, T., Garnero, E. J., and Russell, S. A. High-resolution investigation of shear wave anisotropy in D′′ beneath the Cocos Plate. *Geophysical Research Letters*, 31:L07605, 2004. doi: 10.1029/2003GL018902.

Rokosky, J. M., Lay, T., and Garnero, E. J. Small-scale lateral variations in azimuthally anisotropic D′′ structure beneath the Cocos Plate. *Earth and Planetary Science Letters*, 248:411–425, 2006. doi: 10.1016/j.epsl.2006.06.005.

Russo, R., Gallego, A., Comte, D., Mocanu, V., Murdie, R., and VanDecar, J. Source-side shear wave splitting and upper mantle flow in the Chile Ridge subduction Region. *Geology*, 38:707–710, 2010. doi: 10.1130/G30920.1.

Savage, M. K. Seismic anisotropy and mantle deformation: What have we learned from shear wave splitting? *Reviews of Geophysics*, 37:65 – 106, 1999. doi: 10.1016/10.1029/98RG02075.

Scripps Institution of Oceanography. Global Seismograph Network - IRIS/IDA, 1986. doi: 10.7914/SN/II.

Silver, P. G. Seismic Anisotropy beneath the Continents: Probing the Depths of Geology. *Annual Review of Earth and Planetary Sciences*, 24(1):385 – 432, 1996. doi: 10.1146/annurev.earth.24.1.385.

Silver, P. G. and Chan, W. W. Shear wave splitting and subcontinental mantle deformation. *Journal of Geophysical Research: Solid Earth*, 96:16429–16454, 1991. doi: 10.1029/91JB00899.

Silver, P. G. and Long, M. D. The non-commutivity of shear wave splitting operators at low frequencies and implications for anisotropy tomography. *Geophysical Journal International*, 184 (3):1415–1427, 2011. doi: 10.1111/j.1365-246X.2010.04927.x.

Silver, P. G. and Savage, M. K. The Interpretation of Shear-Wave Splitting Parameters In the Presence of Two Anisotropic Layers. *Geophysical Journal International*, 119:949–963, 1994.

Suzuki, Y., Kawai, K., and Geller, R. J. Imaging paleoslabs and inferring the Clapeyron slope in D′′ beneath the northern Pacific based on high-resolution inversion of seismic waveforms for 3-D transversely isotropic structure. *Physics of the Earth and Planetary Interiors*, page 106751, 2021. doi: 10.1016/j.pepi.2021.106751.

Tesoniero, A., Leng, K., Long, M. D., and Nissen-Meyer, T. Full wave sensitivity of SK(K)S phases to arbitrary anisotropy in the upper and lower mantle. *Geophysical Journal International*, 222(1):412 – 435, 2020. doi: 10.1093/gji/ggaa171.

Thomas, C. and Kendall, J.-M. The lowermost mantle beneath northern Asia—II. Evidence for lower-mantle anisotropy. *Geophysical Journal International*, 151:296–308, 2002. doi: 10.1046/j.1365-246X.2002.01760.x.

Utrecht University (UU Netherlands). NARS, 1983. doi: 110.7914/SN/NR.

Walker, A. and Wookey, J. MSAT - a new toolkit for the analysis of elastic and seismic anisotropy. *Computers and Geosciences*, 49: 81–90, 2012. doi: 10.1016/j.cageo.2012.05.031.

Walpole, J., Wookey, J., Masters, G., and Kendall, J. M. A uniformly processed data set of SKS shear wave splitting measurements: A global investigation of upper mantle anisotropy beneath seismic stations. *Geochemistry, Geophysics, Geosystems*, 15:1991–2010, 2014. doi: 10.1002/2014GC005278.

Walsh, E., Arnold, R., and Savage, M. K. Silver and Chan revisited. *Journal of Geophysical Research: Solid Earth*, 118:5500–5515, 2013. doi: 10.1002/jgrb.50386.

Wang, Y. and Wen, L. Mapping the geometry and geographic distribution of a very low velocity province at the base of the Earth's mantle. *Journal of Geophysical Research: Solid Earth*, 109, 2004. doi: 10.1029/2003JB002674.

Wessel, P. and Smith, W. H. F. New, improved version of generic mapping tools released. *Eos, Transactions American Geophysical Union*, 79:579–579, 1998. doi: 10.1029/98EO00426.

Wolf, J. and Long, M. D. Slab-driven flow at the base of the mantle beneath the northeastern Pacific Ocean. *Earth and Planetary Science Letters*, 594:117758, 2022. doi: 10.1016/j.epsl.2022.117758.

Wolf, J. and Long, M. D. Lowermost mantle structure beneath the central Pacific Ocean: Ultralow velocity zones and seismic anisotropy. *Geochemistry, Geophysics, Geosystems*, 24: e2022GC010853, 2023. doi: 10.1029/2022GC010853.

Wolf, J., Creasy, N., Pisconti, A., Long, M. D., and Thomas, C. An investigation of seismic anisotropy in the lowermost mantle beneath Iceland. *Geophysical Journal International*, 219 (Supplement_1):S152 – S166, 2019. doi: 10.1093/gji/ggz312.

Wolf, J., Long, M. D., Leng, K., and Nissen-Meyer, T. Sensitivity of SK(K)S and ScS phases to heterogeneous anisotropy in the lowermost mantle from global wavefield simulations. *Geophysical Journal International*, 228:366–386, 2022a. doi: 10.1093/gji/ggab347.

Wolf, J., Long, M. D., Leng, K., and Nissen-Meyer, T. Constraining deep mantle anisotropy with shear wave splitting measurements: Challenges and new measurement strategies. *Geophysical Journal International*, 230:507–527, 2022b. doi: 10.1093/gji/ggac055.

Wolf, J., Frost, D. A., Long, M. D., Garnero, E., Aderoju, A. O., Creasy, N., and Bozdağ, E. Observations of Mantle Seismic Anisotropy Using Array Techniques: Shear-Wave Splitting of Beamformed SmKS Phases. *Journal of Geophysical Research: Solid Earth*, 128 (1):e2022JB025556, 2023a. doi: 10.1029/2022JB025556.

Wolf, J., Long, M. D., Creasy, N., and Garnero, E. On the measurement of Sdiff splitting caused by lowermost mantle anisotropy. *Geophysical Journal International*, 2023b. doi: 10.1093/gji/ggac490.

Wolf, J., Long, M. D., Li, M., and Garnero, E. Global Compilation of Deep Mantle Anisotropy Observations and Possible Correlation With Low Velocity Provinces. *Geochemistry, Geophysics, Geosystems*, 24(10):e2023GC011070, 2023c.

Wolf, J., Li, M., Haws, A. A., and Long, M. D. Strong seismic anisotropy due to upwelling flow at the root of the Yellowstone mantle plume. *Geology*, 2024. doi: 10.1130/G51919.1.

Wolf, J., Long, M. D., and Frost, D. A. Ultralow velocity zone and deep mantle flow beneath the Himalayas linked to subducted slab. *Nature Geoscience*, pages 1–7, 2024. doi: 10.1038/s41561-024-01386-5.

Wolfe, C. J. and Silver, P. G. Seismic anisotropy of oceanic upper mantle: Shear wave splitting methodologies and observations. *Journal of Geophysical Research: Solid Earth*, 103(B1):749–771, 1998. doi: 10.1029/97JB02023.

Wookey, J. and Kendall, J.-M. Constraints on lowermost mantle mineralogy and fabric beneath Siberia from seismic anisotropy. *Earth and Planetary Science Letters*, 275:32–42, 2008. doi: 10.1016/j.epsl.2008.07.049.

Wookey, J., Kendall, J.-M., and Rümpker, G. Lowermost mantle anisotropy beneath the north Pacific from differential S-ScS splitting. *Geophysical Journal International*, 161:829–838, 2005. doi: 10.1111/j.1365-246X.2005.02623.x.

Wookey, J., Stackhouse, S., Kendall, J.-M., Brodholt, J., and Price, G. Efficacy of the Post-Perovskite Phase as an Explanation for Lowermost-Mantle Seismic Properties. *Nature*, 438:1004–1007, 2005b. doi: 10.1038/nature04345.

Yuan, K. and Beghein, C. Three-dimensional variations in Love and Rayleigh wave azimuthal anisotropy for the upper 800km of the mantle. *Journal of Geophysical Research: Solid Earth*, 119: 3232–3255, 2014. doi: 10.1002/2013JB010853.

Zhu, H., Yang, J., and Li, X. Azimuthal Anisotropy of the North American Upper Mantle Based on Full Waveform Inversion. *Journal of Geophysical Research: Solid Earth*, 125(2):e2019JB018432, 2020. doi: 10.1029/2019JB018432.

doi:10.26443/seismica.v3i1.1124

SEISMICA

Feasibility of Deep Learning in Shear Wave Splitting analysis using Synthetic-Data Training and Waveform Deconvolution

Megha Chakraborty [1,2], **Georg Rümpker** * [1,2], **Wei Li** [1], **Johannes Faber**[1,3], **Frederik Link** [4], **Nishtha Srivastava** [1,2]

[1]Frankfurt Institute for Advanced Studies, 60438 Frankfurt, Germany, [2]Institute of Geosciences, Goethe University Frankfurt, 60438 Frankfurt, Germany, [3]Institute for Theoretical Physics, Goethe University Frankfurt, 60438 Frankfurt, Germany, [4]The Department of Earth & Planetary Sciences, Yale University, New Haven CT 06511, United States

Author contributions: *Conceptualization*: Georg Rümpker and Megha Chakraborty. *Methodology*: Megha Chakraborty, Georg Rümpker, Johannes Faber and Nishtha Srivastava . *Validation*: Wei Li. *Formal Analysis*: Megha Chakraborty. *Data*: Frederik Link. *Writing - Original draft*: Megha Chakraborty. *Writing - Review & Editing*: Megha Chakraborty, Georg Rümpker, Wei Li, Johannes Faber, Frederik Link and Nishtha Srivastava. *Funding acquisition*: Nishtha Srivastava, Georg Rümpker.

Abstract Teleseismic shear-wave splitting analyses are often performed by reversing the splitting process through the application of frequency- or time-domain operations aimed at minimizing the transverse-component energy of waveforms. These operations yield two splitting parameters, ϕ (fast-axis orientation) and δt (delay time). In this study, we investigate the applicability of a baseline recurrent neural network, SWS-Net, for determining the splitting parameters from pre-selected waveform windows. Due to the scarcity of sufficiently labelled real waveform data, we generate our own synthetic dataset to train the model. The model is capable of determining ϕ and δt with a root mean squared error (RMSE) of 9.7° and 0.14 s on noisy synthetic test data. The application to real data involves a deconvolution step to homogenize the waveforms. When applied to data from the USArray dataset, the results exhibit similar patterns to those found in previous studies with mean absolute differences of 9.6° and 0.16 s in the calculation of ϕ and δt, respectively.

Production Editor:
Gareth Funning
Handling Editor:
Lauren Waszek
Copy & Layout Editor:
Kirsty Bayliss

Signed reviewer(s):
Tom Hudson

Received:
September 27, 2023
Accepted:
January 6, 2024
Published:
March 22, 2024

1 Introduction

The analysis of seismic anisotropy serves as a unique tool for investigating the elusive dynamic processes occurring within the Earth's mantle. Inferring vertically and laterally varying anisotropic structures from surface-recorded seismic waveforms can provide vital constraints for geodynamic models of mantle deformation and flow. The study of teleseismic shear-wave splitting (SWS), a technique in use for over three decades, provides key insights about seismic anisotropy, aiding in the analysis of the dynamic processes within Earth's interior (Long and Silver, 2009; Reiss and Rümpker, 2017; Savage, 1999; Silver and Chan, 1991).

Two primary mechanisms contribute to the development of seismic anisotropy in the Earth's mantle: strain-induced lattice preferred orientation (LPO) of upper mantle minerals such as olivine (resulting from differential motion between the lithosphere and asthenosphere, and mantle flow) (Silver and Chan, 1991) and shape preferred orientation due to the presence of vertically aligned fluid-filled fractures, cracks, and microcracks (Holtzman and Kendall, 2010).

When a shear wave enters an anisotropic medium, it is split into two orthogonally polarized components that propagate at different speeds. This phenomenon can be described by two splitting parameters: the fast axis orientation (the polarization direction of the faster wave) ϕ, and the time delay between the two components δt. While ϕ represents the orientation of the anisotropic materials, δt measures the strength of anisotropy and the extent of the anisotropic material. Teleseismic phases are typically employed to investigate the anisotropic properties of the Earth. The most frequently used phases include SKS, SKKS, and PKS, and are collectively referred to as XKS phases. The conversion of these waves at the core-mantle boundary results in polarization in the direction of the back-azimuth (Jia et al., 2021; Liu and Gao, 2013; Reiss and Rümpker, 2017).

Several software codes have been developed to determine the splitting parameters ϕ and δt through grid search or correlation approaches. Examples of such codes can be found in the works of Silver and Chan (1991) such as Liu and Gao (2013); Savage et al. (2010); Teanby et al. (2004); Wüstefeld et al. (2008); Wuestefeld et al. (2010); Hudson et al. (2023).(Semi-)automatic ap-

*Corresponding author: rumpker@geophysik.uni-frankfurt.de

proaches were recently suggested by Reiss and Rümpker (2017) and Link et al. (2022).

In this paper, we present a baseline model that demonstrates the potential of Deep Learning for the analysis of shear-wave splitting. In a recent study, Zhang and Gao (2022) utilized a Convolutional Neural Network (CNN) for waveform classification to automatically select reliable SWS measurements. However, to the best of our knowledge, a comprehensive analysis to infer anisotropic splitting parameters using deep learning has not yet been presented. Here, we introduce a deep learning model called SWSNet (Shear-Wave Splitting Network) to determine the splitting parameters from pre-selected waveform windows which are used by Link et al. (2022) for their analysis. Due to the lack of sufficient labelled data, the model is trained on synthetic data, simulated under the assumption of a single anisotropic layer (as is the case with traditional methods). A series of deconvolution and reconvolution steps are applied to both the real data and the synthetic data to ensure maximum resemblance. We demonstrate that SWSNet can produce results comparable to those of previous studies such as Liu et al. (2014) when applied to real data from the USArray and obtain mean absolute differences of 9.6° and 0.16 s in the calculation of ϕ and δt, respectively.

The major contributions of this paper can be summarised as follows: (i) to the best of our knowledge this is the first work to explore the applicability of deep learning in determining splitting parameters from waveforms; (ii) as we do not have sufficient labelled real data, we use synthetic data to train our model; (iii) a novel deconvolution and reconvolution approach is applied to remove the source and path effects from the real data to bridge the gap between ideal synthetic waveforms and real waveforms.

2 Methods and Results

For our study we use a supervised learning approach, which is a machine learning paradigm that relies on labelled data for training a model. The Deep Learning model we use learns to map the waveforms to the corresponding labels (in our case ϕ and δt) by minimising the difference between the true and predicted labels defined by the loss function.

In principle, labelled waveform data from shear-wave splitting analyses is available from publications and data archives (see e.g., Barruol et al., 2009). However, for our purposes, the amount of available data is limited, and the labelling may not be as uniform as would be required for efficient training. In order to overcome this limitation, we will use synthetic data as an alternative. Ideally, the generated synthetic waveforms will mimic the properties and characteristics of real data.

2.1 Modeling shear-wave splitting

In our approach, we consider waveform effects due to a single anisotropic layer, which is characterized by a horizontal symmetry axis (referred to as the "fast axis" and oriented at an angle ϕ measured clockwise from North). A vertically incident shear wave splits into horizontally polarized fast and slow components, where the fast component aligns parallel to the symmetry axis, while the slow component is oriented perpendicular to it. Generally, these two quasi-shear waves propagate at different speeds, resulting in a separation by the delay time, δt, as they travel through the layer. A graphical representation of the coordinate systems used is given in Figure S1.

The equations to describe shear-wave splitting in layered structures have recently been summarized by Rümpker et al. (2023). In the frequency domain, the radial and transverse displacement components, after passing through the layer, can be expressed as

$$\begin{pmatrix} u_1^{(r)} \\ u_1^{(t)} \end{pmatrix} = \begin{pmatrix} \cos\theta + i\sin\theta\cos 2\alpha & i\sin\theta\sin 2\alpha \\ i\sin\theta\sin 2\alpha & \cos\theta - i\sin\theta\cos 2\alpha \end{pmatrix} \begin{pmatrix} u_0^{(r)} \\ u_0^{(t)} \end{pmatrix} \quad (1)$$

where $\theta = \omega\delta t/2$, $\alpha = \beta - \phi$ is the angular difference between back-azimuth and fast axis, and index 0 denotes waveforms before passing through the anisotropic layer. For XKS phases in a radially symmetric Earth, we can assume that $u_0^{(t)} = 0$ upon entering the (first) anisotropic layer on the receiver-side leg of the ray path, such that

$$u_1^{(r)} = (\cos\theta + i\sin\theta\cos 2\alpha)\, u_0^{(r)} \quad (2)$$

$$u_1^{(t)} = i\sin\theta\sin 2\alpha\, u_0^{(r)} \quad (3)$$

Note, that for relatively long periods, $T \gg \delta t$ (to first order in θ), this simplifies to

$$u_1^{(r)} \simeq (1 + i\omega\frac{\delta t}{2}\cos 2\alpha)\, u_0^{(r)} \quad (4)$$

$$u_1^{(t)} \simeq i\omega\frac{\delta t}{2}\sin 2\alpha\, u_0^{(r)} \quad (5)$$

where the factor $i\omega$ yields a derivative of the radial-component waveform and the amplitude is modulated by $\sin 2\alpha$. We will use this formulation in the development of our deconvolution approach, to be discussed in subsequent sections.

2.2 Deep Learning Analysis - Synthetic Data

We use synthetic data to train our model. The radial and transverse waveforms are generated with a sampling frequency of 50 Hz for back-azimuths between $0 - 360°$ and fast axis ϕ ranging between $0 - 180°$. Consequently α can vary between $0 - 180°$, since ϕ and $\phi + 180°$ represent the same fast axis orientation. We allow for possible values of δt between 0.2-2.0 seconds. Note that δt characterizes the anisotropy within the layer and is not equal to an "apparent" delay time which could be much larger (e.g. Silver and Savage, 1994). A total of 10^6 waveforms are used for the training process; this dataset is

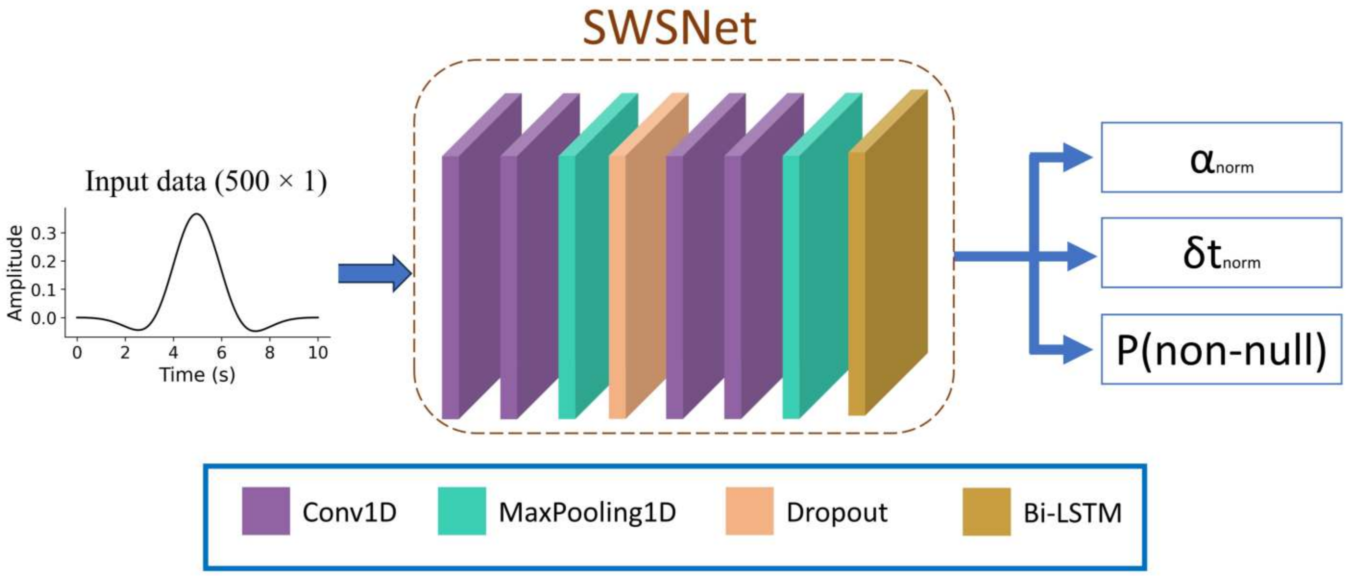


Figure 1 The architecture of SWSNet. The model takes as input the (deconvolved) transverse component and comprises of two blocks of 1D convolution and Maxpooling operations seperated by a Dropout layer with drop rate 30%, and followed by a bi-directional LSTM layers. The final outputs are the normalised values of α (α_{norm}) and δt (δt_{norm}) and the probability of the measurement being non-null.

split in a ratio of 80:20 for training and validation purposes.

Combinations of δt and ϕ are chosen from uniform random distributions for the ranges described above. We experiment with Convolutional layers (Kiranyaz et al., 2015), Bi-directional Long Short-Term Memory (Bi-LSTM) (Hochreiter and Schmidhuber, 1997) layers and a combination of both. Convolutional layers have been established to be effective at feature-extraction, while Bi-LSTMs are known for their ability to detect temporal dependence between these features. The model hyperparameters (such as the number of layers, the kernel size of filters in the convolutional layers, the dimensionality of the LSTM layers, the activation functions to be used etc.) are chosen by experimenting to maximise the model performance on validation data. Each 1D convolutional layer used has a Rectified Linear Unit (ReLU) activation function (Agarap, 2018). The model outputs three values corresponding to the probability of the measurement being non-null and the normalised predictions for δt and ϕ. The normalization of the target variables ensures that the mean squared error loss calculated for them are of the same order; this helps in the convergence of the loss function during backpropagation. Here, any measurement with $\alpha < 2$, $88 < \alpha < 92$, and $\alpha > 178$ is considered a null measurement. Since it is impossible for the model to discriminate between $\alpha = 0°$, $\alpha = 90°$ and $\alpha = 180°$, the transverse component energy for all these cases being zero, we find that defining a non-null class helps the model learn to estimate α. A rectified linear-unit (ReLU) activation function (Agarap, 2018) is used for layers predicting α and δt while a sigmoid function is used to output the probability corresponding the measurement being non-null. A schematic example of such an architecture is shown in Figure 1 (note that Figure 1 shows the final architecture of SWSNet described is section 2.4); a more detailed diagram is provided in the Supplementary information (Figure S7).

The model is trained using the Adam Optimiser (Kingma and Ba, 2014). We use a batch size of 256. Mean squared error and binary cross-entropy are used as loss functions for regression and classification respectively. Apart from using Maxpooling (Nagi et al., 2011) and Dropout (Srivastava et al., 2014) layers in the model architecture, early stopping (Prechelt, 2012) is used to further prevent overfitting, whereby training stops if validation loss does not decrease for 8 consecutive epochs. With this condition the model trains for 35 epochs.

2.2.1 Results - Synthetic Data

We train the Neural Network on a dataset with noise applied independently to the fast and slow components. Two types of noise are experimented with– random and Gaussian. The noise level is chosen from a random normal distribution with mean 30% and standard deviation 10%. Some examples for these datasets can be found in Figures S2 and S3 in the Supplementary Materials. Figure 2 (a) and (b) show the results when the model trained on synthetic data with random noise is tested on an independent test dataset also with random noise. As can be seen from Figure 2, the deep learning model has RMSE 5.9° and 0.12 s in the predictions of α and δt respectively. The corresponding figure (S4) for data with Gaussian noise is provided as Supplementary information.

2.3 Application to real data from USArray

We apply our model to pre-selected waveforms from the USArray dataset and compare our results with Liu et al. (2014) and those calculated by the automatic Splitracer toolbox proposed by Link et al. (2022). To make sure that only meaningful results are used in the calculation

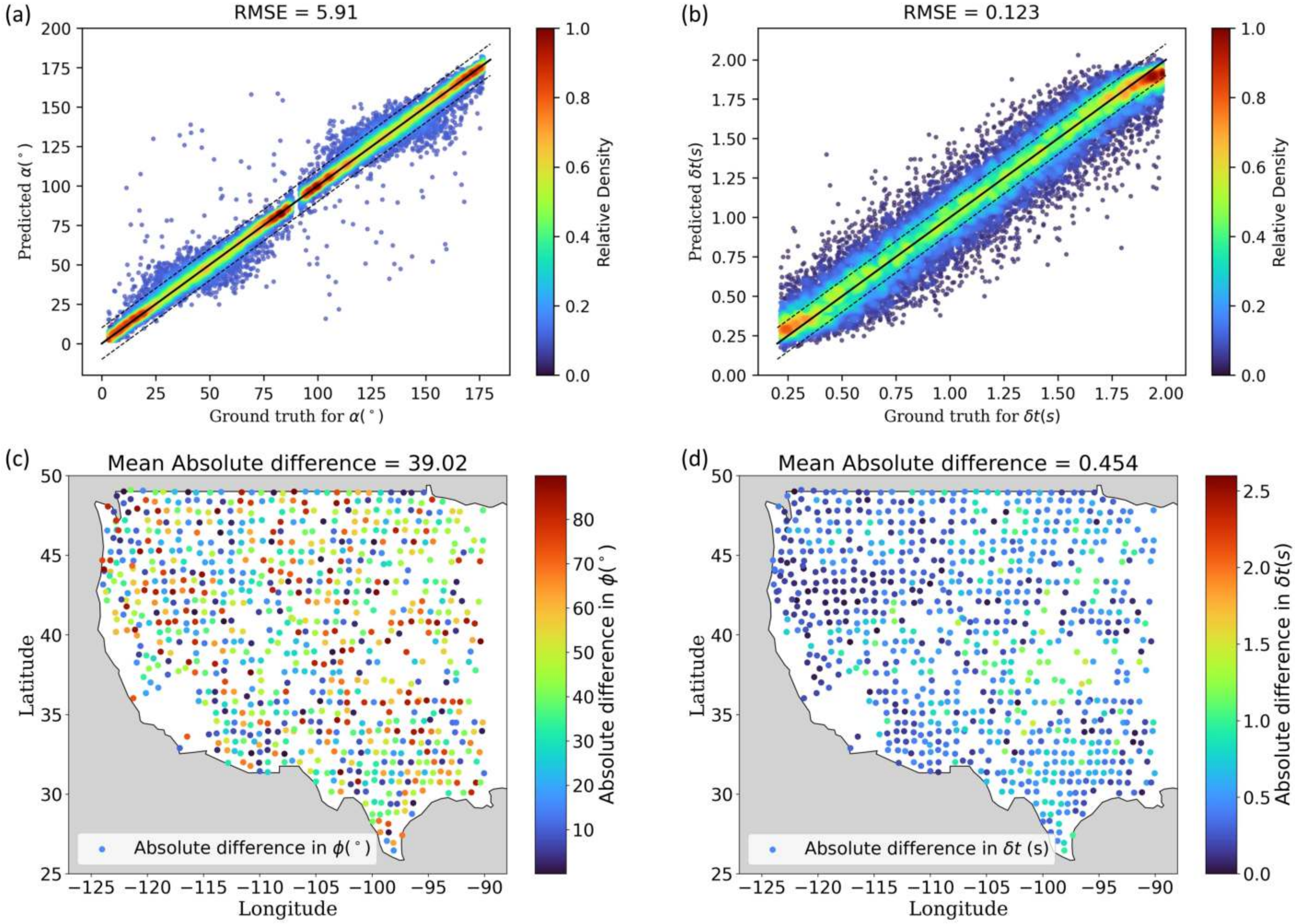


Figure 2 The relation between ground truth and predictions for (a) α and (b) δt when the model trained on synthetic training data contaminated with random noise is tested on synthetic test data contaminated by random noise; comparison between station-wise averages of (c) α and (d) δt calculated using the deep learning model and those given by Liu et al. (2014). (The corresponding figure for data contaminated with Gaussian noise can be found in the Supplementary Materials.)

of station averages, we perform a quality check on the estimations made by the neural network on given waveforms. We perform splitting inversion using the splitting parameters predicted by the neural network and check the percentage reduction in the transverse component energy (sum of squared amplitudes) as proposed by Silver and Chan (1991). An experimentally chosen threshold of 60% reduction in transverse component of energy is used to select the waveforms to be used for calculating station-wise averages for splitting parameters.

2.3.1 Direct application of the Model

When the model trained on the synthetic data is directly applied to the real data (radial and transverse components), the station-wise averages obtained for the splitting parameters differ significantly from those presented by Liu et al. (2014), as shown in Figure 2(c) and (d) (and also, S4 (c) and (d)). This happens as real waveforms look significantly different from the synthetic data. Thus a direct application of the trained model to the real waveforms renders unusable results. This necessitates an intermediate step to bridge the gap between the synthetic and real waveforms.

2.4 Deconvolution approach

Observed real waveforms are not only affected by anisotropic layering but may vary significantly due to different source mechanisms (and path effects). This poses a challenge to the training of the deep learning model, as it is not computationally feasible to include all waveform variations that may arise from different source mechanisms and complexities of the medium. Here, we choose a deconvolution approach to mitigate source effects and "homogenize" the waveforms. This method is similar to the one utilized in receiver-function processing, for instance Langston (1979); Owens et al. (1984); Ammon (1991).

We deconvolve both the radial and transverse components by the radial component. In the frequency-domain, in view of eq. (5), the procedure applied to real data can be described as follows:

$$u_*^{(r)} = u_1^{(r)}/u_1^{(r)} \tag{6}$$

$$= 1$$

$$\begin{aligned} u_*^{(t)} &= i\omega \frac{\delta t}{2} \sin 2\alpha\, u_0^{(r)}/u_1^{(r)} \\ &\simeq i\omega \frac{\delta t}{2} \sin 2\alpha \end{aligned} \tag{7}$$

Note that we assumed $u_0^{(r)}/u_1^{(r)} \simeq 1$ in the derivation of eq. (7). This implies that the radial-component waveform is a sufficient representation of the incoming waveform (before it enters the anisotropic layer), which further limits the applicability to waveforms with periods much longer than δt ($T \gg \delta t$). The value of 1 for the radial component in the frequency domain corresponds to a δ-function in the time domain. For the transverse component, the factor $i\omega$ causes a time-domain derivative (of the unsplit waveform) with amplitude modulated by $\sin 2\alpha$. In a second step, the deconvolved components can now be convolved with a reference waveform, such as the normalised derivative of an exponential function (Figure S5, also shown in the radial component of Figure 6 described in Section 2.5), to yield a uniform radial component, and standard transverse component that depends on the two splitting parameters. Figure S6 shows the appearance of the transverse component for different α and δt pairs.

For real data, first the waveforms within the selected time windows are resampled at 50 Hz and then the mean is removed. For both the real and synthetic data the following steps are applied:

- A Hanning window is applied to smoothen the transition to zero amplitude at the boundaries of the time window.
- The data is zero-padded to have a uniform total of 2000 time samples corresponding to a 40 s time window.
- A butterworth lowpass filter with corner frequency of 1 Hz is applied to suppress higher-frequency noise.
- The radial component is deconvolved from both the radial and transverse components as per equations 6 and 7.
- The clean waveform shown in Figure S6 (also shown in the radial component of Figure 6 described in Section 2.5) is convolved with both the deconvolved waveforms (radial and transverse components).
- A Hanning window is applied to reduce the effect of possible sinusoidal “ringing” on the transverse component of the reconvolved data.
- The waveform is cropped to the central 10 seconds.
- Another Hanning window is applied followed by the normalisation of the data such that the absolute maximum amplitude in the transverse component is 1.

Figure 3 demonstrates the effectiveness of this method in uniforming the waveforms: while the two waveforms with very close splitting parameters look significantly different due to different source mechanism and path effects, upon applying the deconvolution and reconvolution method described above, they look almost the same.

With this approach it is only the transverse component that carries meaningful information about the splitting parameters. Therefore we retrain our model on the transverse component of the de/reconvolved synthetic waveforms. Once again we experiment with different model architectures and hyperparameters; we find the best performing model to be the one shown in Figure 1. This model will henceforth be called the SWSNet (shear-wave splitting network). A detailed description of the hyperparameters used can be seen in Figure S7. As the input data structure is relatively simple a deeper network does not improve the results and a simple network is sufficient. Please note that the labels corresponding to α and δt are always scaled to be in the range 0-1 as this is known to benefit learning. A training data size of 10^6 waveforms is experimentally found to be optimum (Figure S8).

Once again, we experiment with both random and Gaussian noise. The performance of SWSNet on the synthetic dataset can be seen in Figure 4, and the corresponding figure for the Gaussian noise case is shown in Supplementary Materials (Figure S9). It can be noted here that the performance on the synthetic data worsens in comparison to Figure 2. This is because a major difference in the deconvolution approach, as compared to the method discussed in Section 2.2, is that when we train the model on the deconvolved data, only the transverse component carries the relevant information. Hence, the model is trained only on this component, as opposed to the previous method where both the radial and transverse components were used. Using two components might help the model learn the noise characteristics in the data, resulting in a smaller spread in the predicted parameters. However, despite this deterioration in performance on the synthetic test data, the use of the deconvolution method leads to much better generalizability when applied to real-world data, as will be seen in the subsequent discussion and in Figure 7.

2.5 Application to USArray

We apply our final SWSNet to the real data from USArray. The method used to find the station-wise averages is the same as described in Section 2.3; we experiment with the threshold for energy reduction once again, to choose the optimum threshold for our calculations. A threshold of 60% is determined to be optimum based on our observation in Figure 5 as it results in relatively lower mean absolute differences in the station-wise averages of both splitting parameters, while still retaining a good number of stations.

This leaves us with 8699 acceptable waveforms out of a total of 106323 ($\simeq 8.2\%$). This number is very similar to the 7.6% waveforms marked as ‘good’ category by Link et al. (2022). Some examples of SWSNet’s performance on individual waveforms can be seen in Figure 6 and the corresponding splitting parameters are summarised in

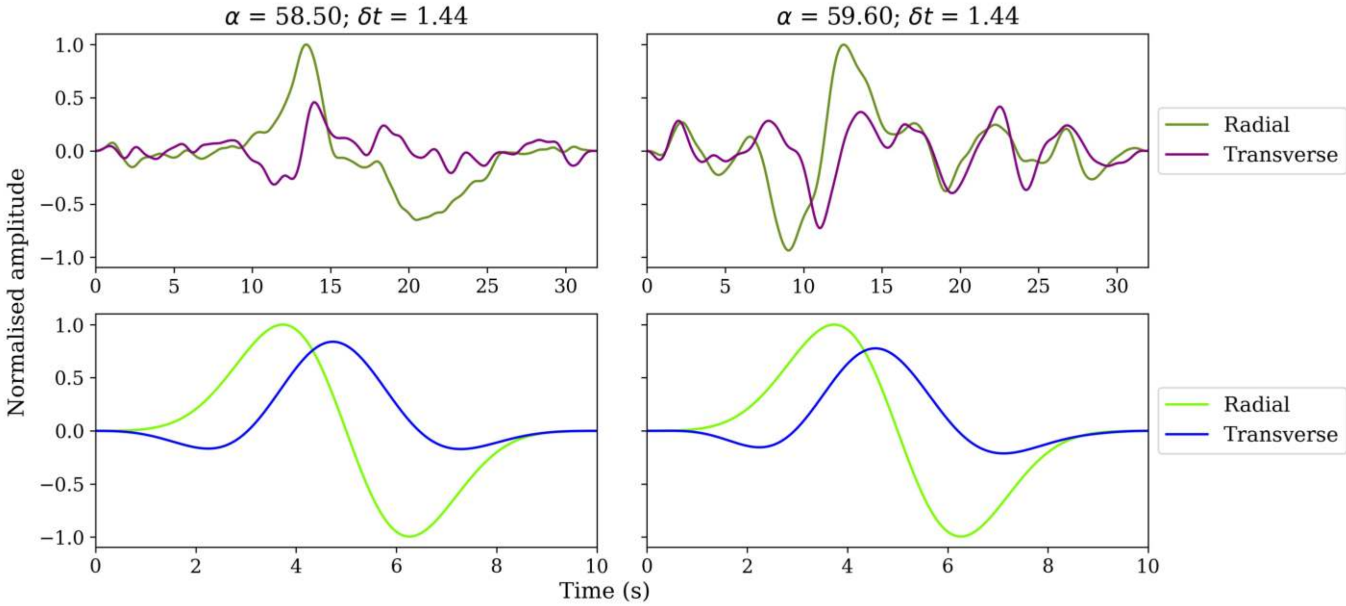


Figure 3 Two example waveforms with splitting parameters calculated by Link et al. (2022) very close to each other (top panels). The bottom panels show the corresponding waveforms after undergoing the deconvolution and reconvolution process described in Section 2.4. While the waveforms in their original form look significantly different due to their respective source mechanism, deconvolution makes them look almost the same, thereby eliminating the source and path effects.

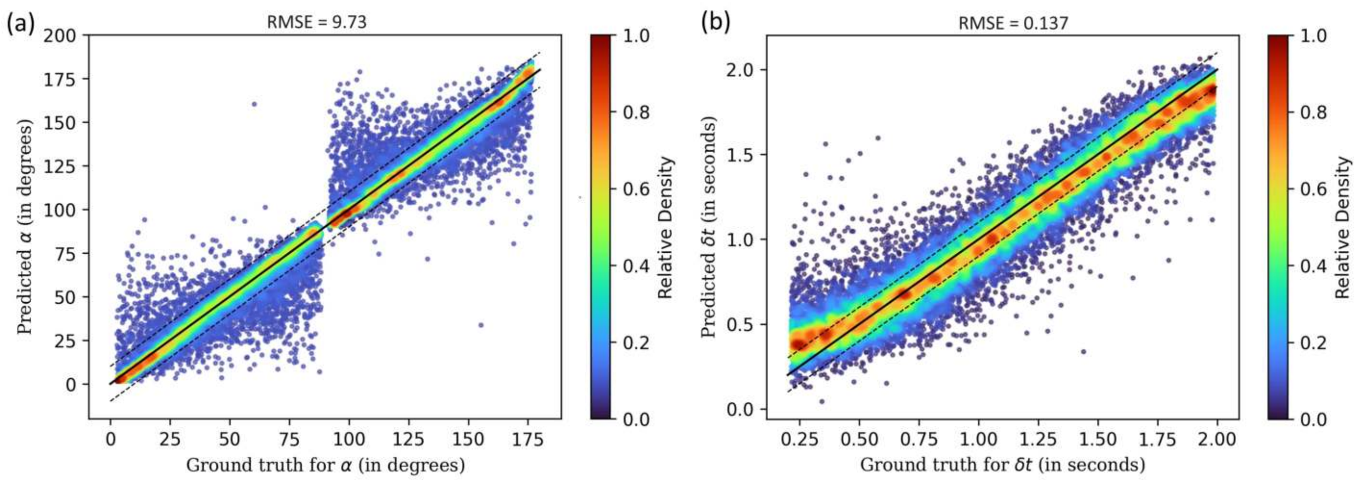


Figure 4 The performance of SWSNet on the synthetic test dataset when including the deconvolution approach. Both the training and test datasets are contaminated by random noise, with noise level chosen from a random normal distribution with mean 30% and standard deviation 10%.

Table 1. One can see the similarity between parameters calculated by SplitRacer, used in Link et al. (2022), and those calculated by SWSNet. A more detailed comparison with grid search results is included in table S1.

Figure 7 shows a visual representation of the station-averages of the splitting parameters calculated by SWSNet and Liu et al. (2014). Unlike Link et al. (2022), Liu et al. (2014) does not employ a joint splitting analysis, allowing for a more direct comparison with our approach, as it is also based on averaging results from individual split phases at a given station. Please note that this model is trained on data with random noise. The results for a model trained on Gaussian noise can be seen in Figure S10. While the performance of the models trained on random and Gaussian noise have comparable performance on the corresponding synthetic test dataset, we observe throughout our experiments that the models trained on data with random noise fit the real data better. We suspect that this is because it is easier for the model to overfit the data with Gaussian noise during training as compared to when the noise is completely random. We also show the comparison between SWSNet and Link et al. (2022) in Figure S11.

3 Discussion

We apply the transverse energy reduction thresholds to SWSNet calculations when calculating station averages. This results in different sets of waveforms being used by this study and by Liu et al. (2014) for these calculations. However, in cases of multi-layer anisotropy, there is a strong dependence of splitting parameters on back-

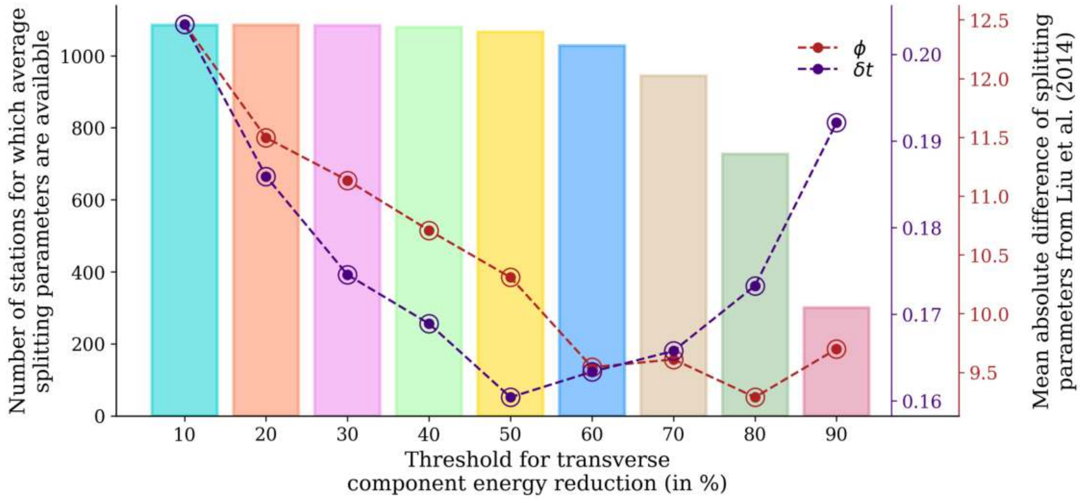


Figure 5 The effect of using different thresholds for energy reduction in the transverse component energy on the final calculation of the station averages. Based on these observations, a threshold of 60% is determined to be optimum. It results in relatively lower mean absolute differences in the station-wise averages of both splitting parameters, while still retaining a good number of stations.

Table 1 A comparison between splitting parameters for individual waveforms shown in Figure 6, calculated by Link et al. (2022) and SWSNet. A detailed comparison between grid search results, results from Link et al. (2022) and SWSNet can be found in table S1 in the Supplementary Information.

Event ID	$\phi(°)$ (Link et al., 2022)	$\phi(°)$ (SWSNet)	δt (s) (Link et al., 2022)	δt (s) (SWSNet)
Y13A2008-05-09T22:15:04SKS	45	49.1	1.33	1.11
P59A2014-08-18T02:55:43SKS	86	88.4	0.82	0.80
121A2018-07-13T10:10:08SKS	9	12.7	1.02	1.03
D25K2017-07-15T12:35:42SKKS	66	71.5	1.44	1.50

azimuth, observed at many locations in the Western/-Central U.S, which could significantly affect the splitting analysis if the events included are not identical.. As such, efforts were made to keep our method free of any requirements of prior knowledge; hence, the threshold was applied for the selection of waveforms irrespective of whether they were used in the calculations by Liu et al. (2014). To understand what the comparison would look like when using the same waveforms in both calculations, we examined a subset of waveforms included in both the station-average calculations by Liu et al. (2014) and in the data used for SWSNet calculations. We recalculated the station averages using just this data and conducted a comparison similar to Figure 5. This figure has been added to the supplementary materials as Figure S12. As expected, we found a closer alignment of the station averages in this case. Furthermore, we compared the splitting parameters calculated by SWSNet and those published by Liu et al. (2014) for individual waveforms, finding that the mean absolute difference for ϕ and δt are 11.08° and 0.239 s respectively.

As a further step to evaluate SWSNet's performance in a multi-layer anisotropy case, we tested it on synthetic waveform data generated by considering two layers of anisotropy with the following two sets of splitting parameters:

1. $\phi_1 = 20°$, $\delta t_1 = 1.0$ s and $\phi_2 = 70°$, $\delta t_2 = 1.0$ s

2. $\phi_1 = 20°$, $\delta t_1 = 1.5$ s and $\phi_2 = 110°$, $\delta t_2 = 0.5$ s

where ϕ_1 and δt_1 represent the fast-axis orientation and time-delay in the first (lower) anisotropic layer, respectively, and ϕ_2 and δt_2 represent the fast-axis orientation and time-delay in the second (upper) anisotropic layer, respectively. Note that in the second case, the fast axes are perpendicular, such that the model effectively corresponds to a model with a single anisotropic layer. The resulting effective delay time is given by the difference between the delay times in each layer. We compare the variation of the splitting parameters with back-azimuth to the theoretical curves calculated as per Silver and Savage (1994) (Figure 8). We find a good agreement between the expected apparent splitting parameters and those predicted by SWSNet, except when the resulting transverse components are very small. These small components correspond to null measurements (indicated by the gray patches in Figure 8) and are attributed by SWSNet to small (case 1) or variable (case 2) delay times. It is further interesting to note that the largest delay times predicted by SWSNet (2 s) agree with maximum delay times used in the training data for a single layer.

We further explore the different factors that affect the station-averaged results, and find the predominant factor to be the number of acceptable measurements for a given station, whereby the difference between the station averaged splitting parameters calculated by SWSNet and those from Liu et al. (2014) diminishes with an increased number of acceptable measurements corre-

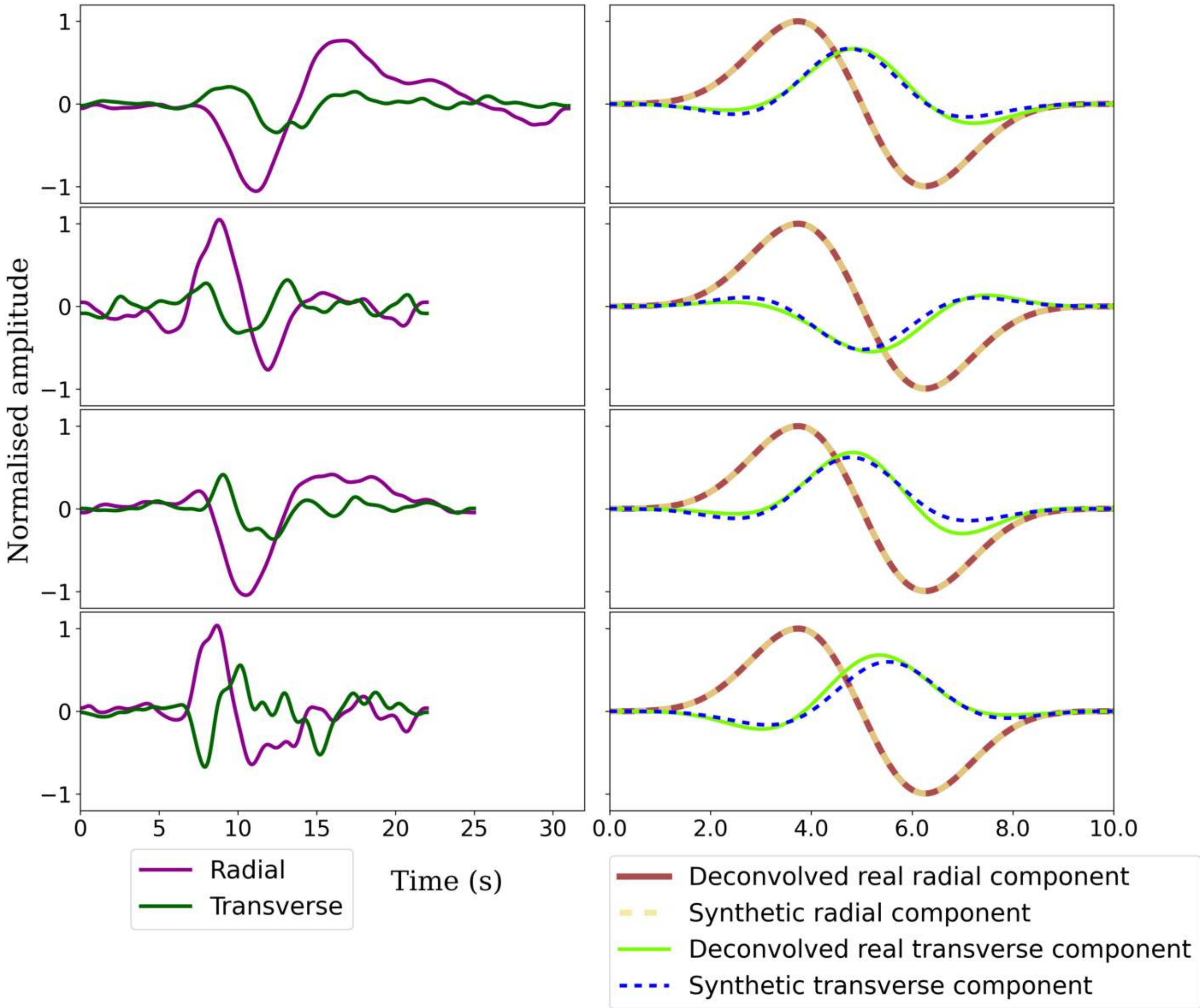


Figure 6 Examples of applying SWSNet to deconvolved real waveforms from the USArray Dataset. The left panel displays the original radial and transverse waveforms. The right panel shows a comparison between the deconvolved real waveforms and the synthetic counterparts, which are generated using the splitting parameters as predicted by SWSNet. The comparison reveals that the radial components are identical, as expected, while the transverse components exhibit a high degree of similarity. The corresponding splitting parameters can be found in Table 1.

sponding to a station (Figure S13).

We also compare our method against a simple grid search algorithm that, like previous studies, finds the splitting parameters for which (upon waveform inversion) the energy in the transverse component is the lowest. The grid search is done between 0.2-2 seconds for δt and 0-180° for α, with a grid spacing of 0.1 second and 1°, respectively. We plot the energy distributions for different combinations of α and δt for five randomly chosen events from five different stations, and find the parameters calculated by SWSNet to be quite close to those found by grid search and what is calculated by Link et al. (2022) (Figure S14). We further observe that grid search on average takes 3-6 times the amount of time taken by SWSNet to calculate splitting parameters for a single waveform.

4 Conclusion

In this study we introduce a baseline deep learning model SWSNet that has the potential to replace grid search methods used by previous studies to find splitting parameters for a waveform. Due to the dearth of labelled real data we train the model on synthetic data. We demonstrate that a direct application of model trained on the synthetic waveforms to real waveforms does not work well, the real waveform being affected by source mechanisms and path effects. This is resolved by using a deconvolution approach to minimise the difference between real and synthetic data. We retrain the model on transverse components of deconvolved synthetic waveforms contaminated by random noise, and show that the model learns to perform reasonably well in identifying the splitting parameters for such waveforms. We then apply our model to pre-selected wave-

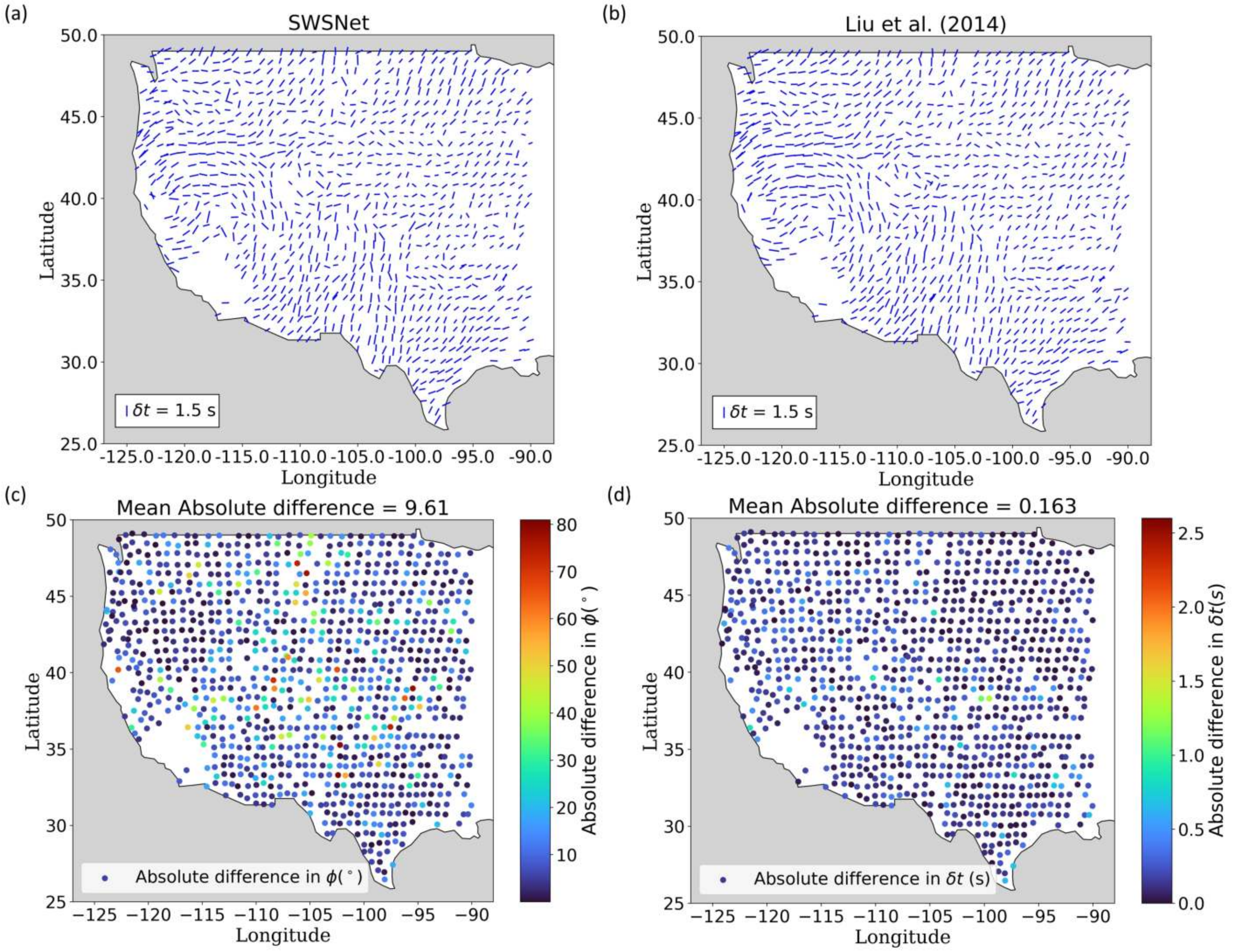


Figure 7 (a) Splitting parameters calculated by SWSNet (b) Splitting parameters calculated by Liu et al. (2014). The orientation of the straight lines is representative of the fast axis orientation while the length represents delay time. A similar general pattern is observed in both cases. (c) Station-wise comparison between ϕ calculated by SWSNet and Liu et al. (2014) (d) Station-wise comparison between δt calculated by SWSNet and Liu et al. (2014)

forms from the USArray dataset and show that the station averages calculated using SWSNet follow the same general trends as previous studies. We observe that the robustness of the proposed method increases with increased number of measurements for a given station.

The current version of the model is trained entirely on synthetic data, but in future versions real data can be added to the training set for improved representation. We would like to reiterate that the approach presented in this work is a baseline method to establish deep learning as a potential candidate for shear wave splitting studies. There are several avenues to further improve the results that would be explored in the future such as using a deeper model or using more complex data, for example, by considering multiple anisotropic layers instead of one. One major draw back of basic neural networks is their inability to provide uncertainty estimates Gawlikowski et al. (2023); therefore, providing uncertainty estimates would be another important avenue to explore in the future.

Acknowledgements

This research is supported by the "KI-Nachwuchswissenschaftlerinnen" - grant SAI 01IS20059 by the Bundesministerium für Bildung und Forschung - BMBF. GR gratefully acknowledges the support provided by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through grants RU 886/15 and RU 886/16. Calculations were performed at the Frankfurt Institute for Advanced Studies' new GPU cluster, funded by BMBF for the project Seismologie und Artifizielle Intelligenz (SAI). The authors are also thankful to Darius Fenner, Jonas Köhler and Kai Zhou for their kind suggestions and to Dr. Tom Hudson and an anonymous reviewer, whose valuable feedback and suggestions have greatly enriched the paper.

Data and code availability

The raw seismic waveforms used in this study are open for download from the IRIS Data Management Center under the network code TA (IRIS Transportable Array,

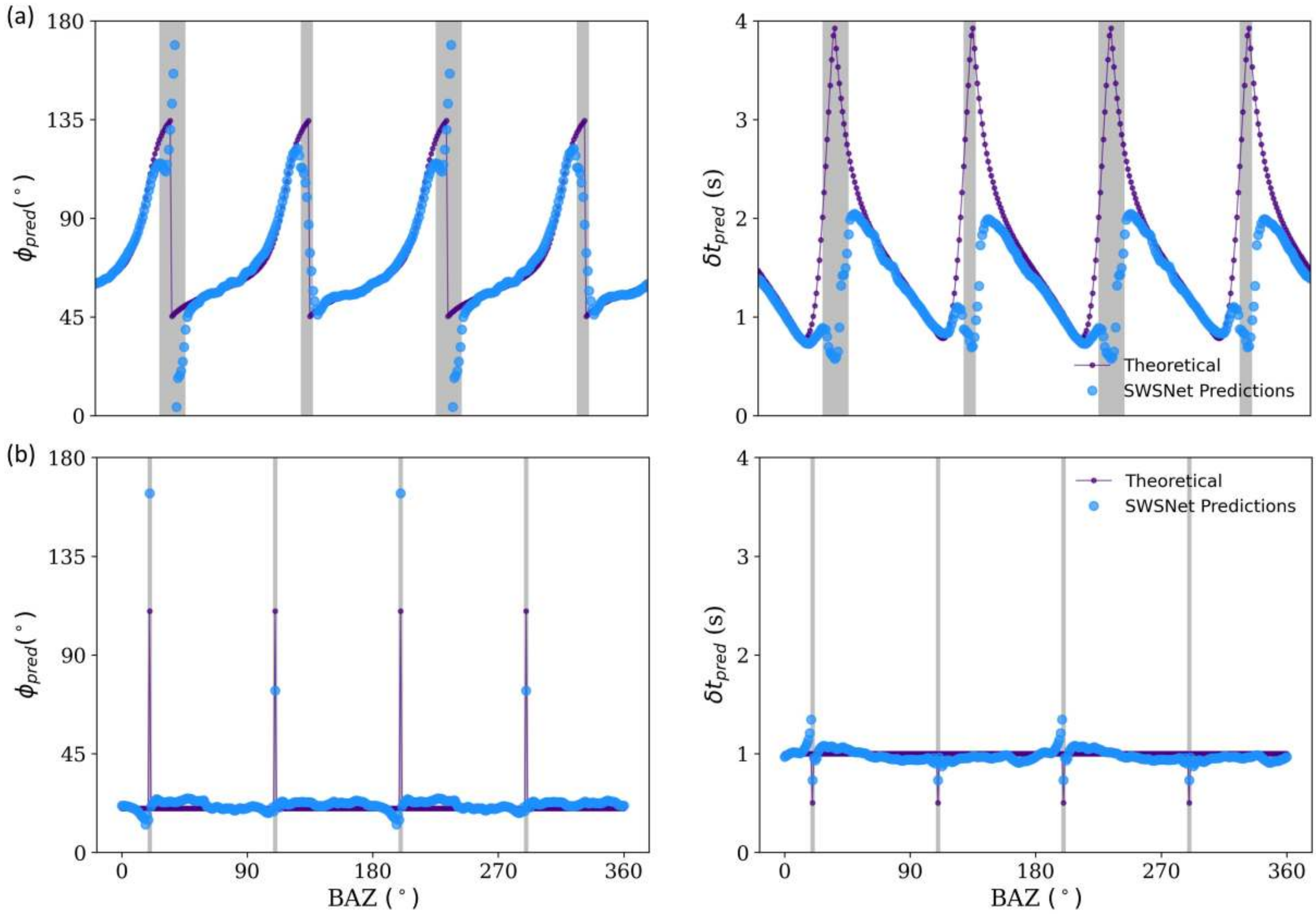


Figure 8 Comparison between apparent splitting parameters calculated according to Silver and Savage (1994) and those predicted by SWSNet for cases of 2-layer anisotropy with the following sets of splitting parameters: (a) $\phi_1 = 20°; \phi_2 = 70°; \delta t_1 = 1.0s; \delta t_2 = 1.0s$ and (b) $\phi_1 = 20°; \phi_2 = 110°; \delta t_1 = 1.5s; \delta t_2 = 0.5s$

2003). The event selection and corresponding labels used for training of SWSNet are available in the supplementary data alongside Link et al. (2022). The codes used for this study can be found at https://github.com/srivastavaresearchgroup/SWSNet.

Competing interests

The authors declare no competing interests.

References

Agarap, A. F. Deep learning using rectified linear units (relu). *arXiv preprint arXiv:1803.08375*, 2018.

Ammon, C. J. The isolation of receiver effects from teleseismic P waveforms. *Bulletin of the Seismological Society of America*, 81 (6):2504–2510, 1991. doi: 10.1785/BSSA0810062504.

Barruol, G., Wuestefeld, A., and Bokelmann, G. SKS-Splitting-database. *Université de Montpellier, Laboratoire Géosciences*, 2009. doi: 10.18715/sks_splitting_database.

Gawlikowski, J., Tassi, C. R. N., Ali, M., Lee, J., Humt, M., Feng, J., Kruspe, A., Triebel, R., Jung, P., Roscher, R., Shahzad, M., Yang, W., Bamler, R., and Zhu, X. X. A survey of uncertainty in deep neural networks. *Artificial Intelligence Reviews*, 56 (Suppl 1):1513–1589, 2023. doi: https://doi.org/10.1007/s10462-023-10562-9.

Hochreiter, S. and Schmidhuber, J. Long short-term memory. *Neural Comput.*, 9(8):1735–1780, 1997. doi: 10.1162/neco.1997.9.8.1735.

Holtzman, B. K. and Kendall, J.-M. Organized melt, seismic anisotropy, and plate boundary lubrication. *Geochemistry, Geophysics, Geosystems*, 11(12), 2010. doi: https://doi.org/10.1029/2010GC003296.

Hudson, T. S., Asplet, J., and Walker, A. M. Automated shear-wave splitting analysis for single- and multi- layer anisotropic media. *Seismica*, 2023. doi: https://doi.org/10.26443/seismica.v2i2.1031.

IRIS Transportable Array. USArray Transportable Array, 2003. doi: 10.7914/SN/TA.

Jia, Y., Liu, K. H., Kong, F., Liu, L., and Gao, S. S. A systematic investigation of piercing-point-dependent seismic azimuthal anisotropy. *Geophysical Journal International*, 227(3): 1496–1511, 07 2021. doi: 10.1093/gji/ggab285.

Kingma, D. P. and Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Kiranyaz, S., Ince, T., Hamila, R., and Gabbouj, M. Convolutional Neural Networks for patient-specific ECG classification. pages 2608–2611, 2015. doi: 10.1109/EMBC.2015.7318926.

Langston, C. A. Structure under Mount Rainier, Washington, inferred from teleseismic body waves. *Journal of Geophysical Research: Solid Earth*, 84(B9):4749–4762, 1979. doi: https://doi.org/10.1029/JB084iB09p04749.

Link, F., Reiss, M. C., and Rümpker, G. An automatized XKS-splitting procedure for large data sets: Extension package for SplitRacer and application to the USArray. *Computers & Geosciences*, 158:104961, 2022. doi: https://doi.org/10.1016/j.cageo.2021.104961.

Liu, K. H. and Gao, S. S. Making Reliable Shear-Wave Splitting Measurements. *Bulletin of the Seismological Society of America*, 103 (5):2680–2693, 10 2013. doi: 10.1785/0120120355.

Liu, K. H., Elsheikh, A., Lemnifi, A., Purevsuren, U., Ray, M., Refayee, H., Yang, B. B., Yu, Y., and Gao, S. S. A uniform database of teleseismic shear wave splitting measurements for the western and central United States. *Geochemistry, Geophysics, Geosystems*, 15(5):2075–2085, 2014. doi: https://doi.org/10.1002/2014GC005267.

Long, M. D. and Silver, P. G. Shear Wave Splitting and Mantle Anisotropy: Measurements, Interpretations, and New Directions. *Surveys in Geophysics*, 30:407–461, 2009. doi: https://doi.org/10.1007/s10712-009-9075-1.

Nagi, J., Ducatelle, F., Di Caro, G. A., Cireşan, D., Meier, U., Giusti, A., Nagi, F., Schmidhuber, J., and Gambardella, L. M. Max-pooling convolutional neural networks for vision-based hand gesture recognition. In *2011 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, pages 342–347, 2011. doi: 10.1109/ICSIPA.2011.6144164.

Owens, T. J., Zandt, G., and Taylor, S. R. Seismic evidence for an ancient rift beneath the cumberland plateau, Tennessee: A detailed analysis of broadband teleseismic P waveforms. *J. Geophys. Res.; (United States)*, 9 1984. doi: 10.1029/JB089iB09p07783.

Prechelt, L. *Early Stopping — But When?*, pages 53–67. Springer Berlin Heidelberg, 2012. doi: 10.1007/978-3-642-35289-8_5.

Reiss, M. C. and Rümpker, G. SplitRacer: MATLAB Code and GUI for Semiautomated Analysis and Interpretation of Teleseismic Shear-Wave Splitting. *Seismological Research Letters*, 88(2A): 392–409, 01 2017. doi: 10.1785/0220160191.

Rümpker, G., Kaviani, A., Link, F., Reiss, M., and Komeazi, A. Testing observables for teleseismic shear-wave splitting inversions: ambiguities of intensities, parameters, and waveforms. *Ann. Geophys.*, 66, 2023. doi: https://doi.org/10.4401/ag-8870.

Savage, M. K. Seismic anisotropy and mantle deformation: What have we learned from shear wave splitting? *Reviews of Geophysics*, 37(1):65–106, 1999. doi: https://doi.org/10.1029/98RG02075.

Savage, M. K., Wessel, A., Teanby, N. A., and Hurst, A. W. Automatic measurement of shear wave splitting and applications to time varying anisotropy at Mount Ruapehu volcano, New Zealand. *Journal of Geophysical Research: Solid Earth*, 115(B12), 2010. doi: https://doi.org/10.1029/2010JB007722.

Silver, P. G. and Chan, W. W. Shear wave splitting and subcontinental mantle deformation. *Journal of Geophysical Research: Solid Earth*, 96(B10):16429–16454, 1991. doi: https://doi.org/10.1029/91JB00899.

Silver, P. G. and Savage, M. K. The Interpretation of Shear-Wave Splitting Parameters In the Presence of Two Anisotropic Layers. *Geophysical Journal International*, 119(3):949–963, 12 1994. doi: 10.1111/j.1365-246X.1994.tb04027.x.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15(56):1929–1958, 2014. http://jmlr.org/papers/v15/srivastava14a.html.

Teanby, N. A., Kendall, J.-M., and van der Baan, M. Automation of Shear-Wave Splitting Measurements using Cluster Analysis. *Bulletin of the Seismological Society of America*, 94(2):453–463, 04 2004. doi: 10.1785/0120030123.

Wuestefeld, A., Al-Harrasi, O., Verdon, J. P., Wookey, J., and Kendall, J. M. A strategy for automated analysis of passive microseismic data to image seismic anisotropy and fracture characteristics. *Geophysical Prospecting*, 58(5):755–773, 2010. doi: https://doi.org/10.1111/j.1365-2478.2010.00891.x.

Wüstefeld, A., Bokelmann, G., Zaroli, C., and Barruol, G. SplitLab: A shear-wave splitting environment in Matlab. *Computers & Geosciences*, 34(5):515–528, 2008. doi: https://doi.org/10.1016/j.cageo.2007.08.002.

Zhang, Y. and Gao, S. S. Classification of Teleseismic Shear Wave Splitting Measurements: A Convolutional Neural Network Approach. *Geophysical Research Letters*, 49(12):e2021GL097101, 2022. doi: https://doi.org/10.1029/2021GL097101.

doi:10.26443/seismica.v3i1.1119

SEISMICA

Comparison of geodetic slip-deficit and geologic fault slip rates reveals that variability of elastic strain accumulation and release rates on strike-slip faults is controlled by the relative structural complexity of plate-boundary fault systems

J. Gauriau [*1], **J.F. Dolan** [1]

[1]Department of Earth Sciences, University of Southern California, Los Angeles, CA, USA

Author contributions: *Conceptualization*: J. Gauriau, J. F. Dolan. *Data Curation*: J. Gauriau. *Formal Analysis*: J. Gauriau. *Investigation*: J. Gauriau. *Methodology*: J. Gauriau. *Project Administration*: J. Gauriau. *Resources*: J. Gauriau, J. F. Dolan. *Software*: J. Gauriau. *Supervision*: J. F. Dolan. *Validation*: J. F. Dolan. *Visualization*: J. Gauriau. *Writing – original draft*: J. Gauriau. *Writing – review & editing*: J. F. Dolan, J. Gauriau.

Abstract Comparison of geodetic slip-deficit rates with geologic fault slip rates on major strike-slip faults reveals marked differences in patterns of elastic strain accumulation on tectonically isolated faults relative to faults that are embedded within more complex plate-boundary fault systems. Specifically, we show that faults that extend through tectonically complex systems characterized by multiple, mechanically complementary faults (that is, different faults that are all accommodating the same deformation field), which we refer to as high-Coefficient of Complexity (or high-CoCo) faults, exhibit ratios between geodetic and geologic rates that vary and that depend on the displacement scales over which the geologic slip rates are averaged. This indicates that elastic strain accumulation rates on these faults change significantly through time, which in turn suggests that the rates of ductile shear beneath the seismogenic portion of faults also vary through time. This is consistent with models in which mechanically complementary faults trade off slip in time and space in response to varying mechanical and stress conditions on the different component faults. In marked contrast, structurally isolated (or low-CoCo) faults exhibit geologic slip rates that are similar to geodetic slip-deficit rates, regardless of the displacement and time scales over which the slip rates are averaged. Such faults experience relatively constant geologic fault slip rates as well as constant strain accumulation rate (aside from brief, rapid post-seismic intervals). This suggests that low-CoCo faults "keep up" with the rate imposed by the relative plate-boundary condition, since they are the only structures in their respective plate-boundary zone that can effectively accommodate the imposed steady plate motion. We hypothesize that the discrepancies between the small-displacement average geologic slip rates and geodetic slip-deficit rates may provide a means of assessing a switch of modes for some high-CoCo faults, transitioning from a slow mode to a faster mode, or vice versa. If so, the differences between geologic slip rates and geodetic slip-deficit rates on high-CoCo faults may indicate changes in a fault's behavior that could be used to refine next-generation probabilistic seismic hazard assessments.

Non-technical summary Geodetic slip-deficit rates record how much elastic strain energy accumulates along a fault, whereas geologic slip rates record the actual slip that occurred throughout multiple earthquakes along that fault during release of the stored elastic strain energy. We look at multiple active faults within strike-slip plate boundary fault systems, and compare geodetic slip-deficit rates with geologic slip rates averaged over different fault displacement scales. We find that these values tend to be similar for isolated faults at all scales, whereas they differ in structurally complex fault systems. We conclude that both the accumulation and release of elastic strain energy is constant on faults embedded in simple settings, but varies in complex fault systems.

Production Editor:
Gareth Funning
Handling Editor:
Ake Fagereng
Copy & Layout Editor:
Hannah F. Mark

Signed reviewer(s):
Zoë Mildon
Sam Wimpenny

Received:
September 20, 2023
Accepted:
January 19, 2024
Published:
tbd

1 Introduction

Unravelling the relationship between geologic fault slip rates and rates of strain accumulation as measured by geodesy is critically important for developing a better understanding of the mechanics of faults and the seismic hazards that they pose. Whereas some major faults exhibit constant behavior, with relatively steady geologic slip rates spanning a range of time and displacement scales (e.g., Kozacı et al., 2009, 2011; Berryman et al., 2012; Salisbury et al., 2018; Grant Ludwig et al., 2019), other faults exhibit highly irregular slip rates through time, with centennial to millennial periods of relatively fast slip rate spanning multiple earthquake cycles, separated by prolonged periods of slower or no slip rate (e.g., Benedetti et al., 2002; Friedrich et al., 2003;

*Corresponding author: gauriau@usc.edu

Bull et al., 2006; Dolan et al., 2016; Hatem et al., 2020; Zinke et al., 2017, 2019, 2021).

Elastic strain accumulation rates inferred from analysis of geodetic data reflect the shearing velocity of the seismogenic faults' underlying ductile roots, and have been suggested to be relatively constant beyond the single-earthquake scale (i.e., once fast post-seismic and slower interseismic rates have been averaged out). Indeed, comparisons of geodetic slip-deficit and geologic rates have been used to infer near-constant interseismic rates. For example, in one of the largest such compilations to date, Meade et al. (2013) compared geologic fault slip rates and geodetic slip-deficit rates for 15 major continental strike-slip faults around the world. Their results suggest that, as an ensemble, these faults exhibit a near 1:1 relationship (with a slope of 0.94 ± 0.09) between geologic and geodetic rates. Slight differences between the datasets could be attributable to short-lived periods of higher-than-average strain accumulation during the post-seismic period. The geologic rates used as inputs into the analysis of Meade et al. (2013) span a huge range of displacement and time scales, from as small as ~13 m to as large as ~600 m, and as short as 2 ky to as long as 160 ky. We recently presented results that demonstrate that, for faults that lie within complex plate-boundary fault networks, geologic slip rates vary depending on the displacement scale over which the slip rate is estimated; on the other hand, structurally isolated faults that accommodate most of the relative motion within simple plate boundaries exhibit steadier slip rates (Gauriau and Dolan, 2021). These observations lead us to explore the possibility that differences between geodetic slip-deficit rates and geologic slip rates might also be sensitive to the relative complexity of the surrounding fault network. If they are, this would require that geodetic-geologic rate comparisons consider time and displacement scales over which incremental slip rates are averaged, as well as the relative structural complexity of the surrounding fault system, especially in structurally complex plate boundaries (e.g., northern and southern California, Marlborough fault system in New Zealand), that are characterized by multiple, mechanically complementary faults.

In this paper, we explore the potential constancy, or lack thereof, of the elastic strain accumulation rate patterns on active strike-slip faults. Specifically, we aim to investigate the relative constancy and potential variability of elastic strain accumulation rates on faults characterized by temporally constant geologic slip rates, on the one hand, and faults that exhibit temporally variable geologic slip rates, on the other. Comparing elastic strain accumulation rates derived from geodesy with geologic slip rates has been done in several studies (e.g., Kozacı et al., 2009; Meade et al., 2013; Tong et al., 2014; Dolan and Meade, 2017; Evans et al., 2016) but never in light of the relative complexity of the plate-boundary fault systems being considered.

2 Studied faults and terminology

In this study, we use the recently developed Coefficient of Complexity (CoCo) method (Gauriau and Dolan, 2021), which quantifies the relative structural complexity of the fault network surrounding a fault of interest by integrating the density and displacement rates of the faults in the plate-boundary network at a specific radius (here, 100 km) around the site of interest. The method is illustrated in Figure 1. We use CoCo values calculated for 18 major strike-slip faults for which both geologic incremental slip-rate records and geodetic slip-deficit rates are available (Figure 2, Table 1). In total, we work with 24 different fault sites where these two kinds of data are available and approximately collocated. The comparison of the CoCo values for all sites is then enabled by the standardization of the CoCo values by the respective plate-motion rate, totaled for the observation area of 100 km radius. This allows direct comparisons of the intensity of fault activity in different plate-boundary fault networks that move at different relative plate motion rates.

We divide the available geologic slip-rate data into two groups: large-displacement slip rates and small-displacement slip rates (usually referred to as "long-term" and "short-term" slip rates, respectively), which are averaged over large (> ~50 m) and small (< ~50 m) displacements, respectively (Table 1). The reasons for this are twofold: (a) This allows us to discuss fast- and slow-slipping faults with comparable parameters and hence by considering similar numbers of earthquakes on faults that have widely different recurrence intervals, and (b) displacement, not time, may be what matters most in terms of the mechanisms governing fault behavior in complex plate-boundary fault systems (Dolan et al., 2007; Cawood and Dolan, submitted; Dolan et al., 2024). In addition, we use the terms "geodetic slip-deficit rates" to refer to any rate that was obtained on the basis of space geodetic measurements of surface ground displacement over multi-annual to decadal time scales, such as Global Positioning System (GPS) or Interferometric Synthetic Aperture Radar (InSAR), and which has been modeled to characterize the most recent rate of elastic strain accumulation for the studied strike-slip faults.

3 Consideration of elapsed time since most recent event relative to sampling geodetic slip-deficit rates

In order to evaluate potential differences in behavior of faults embedded within structurally simple fault systems (i.e., low-CoCo faults) versus faults embedded within structurally complex fault networks (i.e., high-CoCo faults), we compare geodetic slip-deficit rates with geologic fault slip rates that are averaged over both small displacements and large displacements. We first introduce a few key considerations that allow us to carry out this comparison between geodetic and geologic data.

The interseismic geodetic data used in this paper may derive from different sampling times throughout the earthquake cycle. Although we have no precise control over where exactly the examined faults lie in their elastic strain cycles, we can in most instances document the elapsed time since their most recent event (MRE),

Fault	Section/ Site	#	SD slip rate (mm/yr)	Time range of SD slip rate (ky)	Displace-ment of SD slip rate (m)	References for SD slip rate	LD slip rate (mm/yr)	Time range of LD slip rate (ky)	Displace-ment of LD slip rate (m)	References for LD slip rate	Geodetic rate (mm/yr)	References for geodetic rate	Plate rate (mm/yr)	References
Garlock	Central	1	$14^{+2.2}_{-1.8}$	1.9	$26^{+3.5}_{-2.5}$	(Dolan et al., 2016)	8.8 ± 1.0	8.0 ± 0.9	70 ± 7	(Fougere et al., 2023)	2.61 ± 3.00	(Evans, 2017b)	49	(Dolan et al., 2016); (McGill and Sieh, 1993; Evans, 2017a)
San Andreas	Mojave	2	$28.8^{+1.5}_{-0.8}$	$1.007^{+0.028}_{-0.050}$	~29	(Weldon et al., 2004; Dolan et al., 2016)	$30.9^{+2.9}_{-2.5}$	1.49 ± 0.13	46	(Weldon et al., 2004)	15.12 ± 2.78		49	(Weldon et al., 2004; Dolan et al., 2016; Evans, 2017a)
	Carrizo Plain	3	$31.6^{+9.0}_{-6.6}$	0.38 ± 0.06	12 ± 1	(Salisbury et al., 2018)	36 ± 1	~3.5	128 ± 1	Grant-Ludwig et al. (2019)	35.65 ± 5.11		39	(Grant Ludwig et al., 2019; Salis-bury et al., 2018; Sieh and Jahns, 1984; Noriega et al., 2006)
San Jacinto	Clare-mont	4	12.8 - 18.3	2.05 ± 0.12	25 - 30	(Onderdonk et al., 2015)					13.18 ± 4.61		49	(Onderdonk et al., 2015; Evans, 2017b)
Owens Valley		5	0.5-2.1 §			(Haddon et al., 2016) and refer-ences therein	2.8-4.5	55-80	235 ± 15	(Kirby et al., 2008)	2.71 ± 1.38		12	(Kirby et al., 2008; Haddon et al., 2016; Evans, 2017a)
Calico	Central	6					1.6 ± 0.2	650 ± 100	900 ± 200	(Oskin et al., 2007)	7.42 ± 3.44		49	(Oskin et al., 2004, 2008; Evans, 2017a)
Hope	Conway	7	$8.2^{+5.4}_{-3.0}$	ca. 1.1	12 ± 2	(Hatem et al., 2020)	$15.2^{+2.2}_{-2.4}$	ca. 13.8	210 ± 15	(Hatem et al., 2020)	$5.8^{+1.8}_{-1.1}$	(Johnson et al., 2022)	39	(Hatem et al., 2020; Johnson et al., 2022)
Wairau	Branch River Dunbeath	8	4.5 ± 1.0 *	3.3 ± 0.4	15 ± 2.6	(Zinke et al., 2021)	4.9 ± 0.4	$11.9^{+1.0}_{-0.8}$	58.5 ± 2	(Zinke et al., 2021)	$2.8^{+2.4}_{-0.8}$			(Zinke et al., 2021; Johnson et al., 2022)
Clarence	Tophouse Road	9	2.0 ± 0.4	$4.5^{+0.8}_{-0.7}$	9.0 ± 1.0	(Zinke et al., 2019)	4.2 ± 0.5	11.2 ± 1.3	47.0 ± 3.0	(Zinke et al., 2019)	$8.6^{+1.4}_{-1.1}$			(Zinke et al., 2019; Johnson et al., 2022)
Awatere	Saxton River	10	$4.2^{+1.2}_{-1.0}$	1.8 ± 0.3	9.5 ± 1.0	(Zinke et al., 2017)	$5.6^{+0.8}_{-0.6}$	$12.9^{+1.2}_{-1.0}$	72.5 ± 7.5	(Zinke et al., 2017)	$1.9^{+2.2}_{-0.8}$			(Zinke et al., 2017; Johnson et al., 2022)
Alpine	Southern	11					$29.6^{+4.5}_{-2.5}$	270	8000	(Barth et al., 2014)	$29.1^{+1.1}_{-3.2}$			(Berryman et al., 2012; Page et al., 2018; Wallace et al., 2012)

Table 1 Continued on next page

Fault	Section/ Site	#	SD slip rate (mm/yr)	Time range of SD slip rate (ky)	Displace-ment of SD slip rate (m)	References for SD slip rate	LD slip rate (mm/yr)	Time range of LD slip rate (ky)	Displace-ment of LD slip rate (m)	References for LD slip rate	Geodetic rate (mm/yr)	References for geodetic rate	Plate rate (mm/yr)	References
Dead Sea	Wadi Araba Valley	12	3.8 - 6.1	2 - 4.2	13.2 ± 1.0	(Klinger et al., 2000)	4 ± 2	140 ± 31	300-900	(Klinger et al., 2000)	5.0 ± 0.2	(Gomez et al., 2020)	7	(Klinger et al., 2000; Niemi et al., 2001; Hamiel et al., 2018)
	Beteiha	13	3.5 ± 0.2 $^{\$}$	1.472	5.2 ± 0.3	(Wechsler et al., 2018)					4.8 ± 0.3			(Wechsler et al., 2018; Masson et al., 2015)
Yam-mouneh		14	3.5 - 7.5	6 - 10	40 ± 5	(Daëron et al., 2004)	2.7-7.3	12 - 27	80 ± 8	(Daëron et al., 2004)	2.5 ± 0.5			(Daëron et al., 2004; Gomez et al., 2003, 2007)
Queen Char-lotte		15					52.9 ± 3.2	17 ± 0.7	900 ± 40	(Brothers et al., 2020)	46.3 ± 0.6	(Elliott and Freymueller, 2020)	55	(Brothers et al., 2020; Elliott and Freymueller, 2020)
Denali	Central	16					12.1 ± 1.7	12.0 ± 1.3 / 11.9 ± 1.3 $^{\#}$	144 ± 14	(Matmon et al., 2006)	7.0 ± 0.3		17	(Matmon et al., 2006; Elliott and Freymueller, 2020; Bender et al., 2023)
	Western	17	10.4 ± 3.0	2.4 ± 0.3	$25\ ^{+5}_{-7}$	(Matmon et al., 2006)	9.4 ± 1.6	16.8 ± 1.8	158 ± 14	(Matmon et al., 2006)	7.75 ± 0.3		17	(Matmon et al., 2006; Elliott and Freymueller, 2020; Bender et al., 2023)

Table 1 Continued on next page

Fault	Section/ Site	#	SD slip rate (mm/yr)	Time range of SD slip rate (ky)	Displacement of SD slip rate (m)	References for SD slip rate	LD slip rate (mm/yr)	Time range of LD slip rate (ky)	Displacement of LD slip rate (m)	References for LD slip rate	Geodetic rate (mm/yr)	References for geodetic rate	Plate rate (mm/yr)	References
Altyn Tagh	Central	18	9.4 ± 0.9 [¤]	5.889 – 5.658	54 ± 5	(Cowgill, 2007)	9.4 ± 2.3	16.6 ± 3.9	156 ± 10	(Cowgill et al., 2009)	9 ± 4	(Bendick et al., 2000)	11.2	(Cowgill, 2007; Cowgill et al., 2009; Bendick et al., 2000; Shen et al., 2001; He et al., 2013; Zhang et al., 2007)
Kunlun	Central Western	19	10.7 ± 2.2	2.885 ± 0.285	31 ± 2	(Haibing et al., 2005)	10.6 ± 1.8	5.96 ± 0.450	63 ± 5	(Haibing et al., 2005)	11.3 ± 3.5	(Zhao et al., 2022)	12	(Van Der Woerd et al., 2002; Haibing et al., 2005; Kirby et al., 2007)
Haiyuan	Laohushan	20	3.7 ± 0.6	9 - 11	32 - 42	(Liu et al., 2022)	4.8 ± 0.2	15 - 17	73 - 79	(Liu et al., 2022)	$5.6^{+1.3}_{-1.1}$	(Daout et al., 2016)	6.5	(Liu et al., 2018; Li et al., 2009; Shao et al., 2020)
North Anatolian	Demir Tepe Eksik	21	16.8 ± 0.1 [*]	0.988	15.3 ± 0.1	(Kondo et al., 2010)	20.5 ± 5.5	2 - 2.5	46 ± 10	(Kozacı et al., 2007)	20.5	(DeVries et al., 2016)	21	(Kozacı et al., 2007; Hubert-Ferrari et al., 2002)
	Tahtaköprü	22					$18.6^{+3.5}_{-3.3}$	~3	55 ± 10	(Kozacı et al., 2009)	21.2 - 21.5		21	(Kozacı et al., 2009)
	Northern / Ganos	23	15 ± 6	2.5 ± 0.5	35.4 ± 1.5	(Meghraoui et al., 2012)	$18.5^{+10.9}_{-5.9}$	490 ± 100	>~8000	(Kurt et al., 2013)	28.6		27	(Meghraoui et al., 2012; Kurt et al., 2013)
East Anatolian	Pazarcık, Tevekkelli	24					5.6 ± 0.3	17.8	101 ± 5	(Yönlü and Karabacak, 2023)	10.3 ± 0.6	(Aktug et al., 2016)	10	(Güvercin et al., 2022; Reilinger et al., 2006)

Table 1 Summary of data from the different fault sections used in this study, including small-displacement (SD), large-displacement (LD) averaged geologic slip rates with corresponding time and displacement ranges over which they are averaged, and geodetic slip-deficit rates. The rate values are reported as they were in their original source publications, unless specified otherwise.

[*] rate calculated between MRE and given offset marker
[§] based on several studies cited in Haddon et al. (2016), with offsets ranging from 3 m (1 earthquake) to 19 m, and respective ages ranging from 600 years ago and 25 ka
[$] averaged over the past four historical earthquakes
[#] first age relates to boulder samples, second age refers to sediment samples (^{10}Be technique)
[¤] using their upper-terrace reconstruction (Cowgill et al., 2009), as for the small-displacement slip rate

as well as an estimate of their mean earthquake recurrence interval. For a majority of the faults we study, it has been at least 100 years since the MRE, as documented historically (e.g., the 1717 Alpine fault earthquake, the 1857 Fort Tejon earthquake on the San Andreas fault, the 1872 Owens Valley earthquake) or on the basis of paleoseismological evidence (e.g., the ca. 1800-1840 CE earthquake on the Conway section of the Hope fault; Hatem et al., 2019). In a few instances, the MRE occurred more recently, such as the series of earthquakes on the North Anatolian fault between 1939 and 1999 (Barka, 1992; Barka et al., 2002), the 2002 Denali earthquake (Haeussler, 2004), or the Kahramanmaraş earthquake (e.g., Barbot et al., 2023) that occurred in Februray 2023 on the East Anatolian fault (for which we use a geodetic rate that was acquired before the earthquake).

Table S1 summarizes the MRE dates and the available mean recurrence intervals for the fault locations we study. In most of the examples, we are well into at least the middle part of the elastic strain accumulation cycle, likely well past any rapid post-seismic deformation (with the possible exceptions of the 1992 Landers, 1999 Izmit, 1999 Düzce, 1999 Hector Mine, and 2002 Denali earthquakes).

4 Relative structural complexity of the surrounding fault network in interpretation of geodetic slip-deficit rate and geologic slip-rate comparisons

In our original formulation of the CoCo metric (Gauriau and Dolan, 2021), we categorized faults as either low- or high-CoCo. To determine the CoCo metric for each fault study site, we apply a system in which we recognize that the degree of structural complexity surrounding a fault is a continuum, with no hard boundary between high- and low-CoCo faults. Whereas many of the faults we study can be readily categorized as either high-CoCo faults (e.g., the Hope fault or the Mojave section of the San Andreas fault) or low-CoCo faults (e.g., the southern Alpine fault, the central San Andreas fault), some of the faults exhibit intermediate CoCo values reflecting a surrounding plate-boundary zone that shows minor to moderate complexity. The two faults that fall in this in-between area are the Central Denali fault (16), characterized by a standardized CoCo value of $1.62 \cdot 10^{-2}$ yr^{-1} and the Altyn Tagh fault (18), characterized by a standardized CoCo value of $1.56 \cdot 10^{-2}$ yr^{-1}. Based on these two values, we use a standardized CoCo value of $1.6 \cdot 10^{-2}$ yr^{-1} as the dividing line between what we will refer to hereafter as low- and high-CoCo faults. With this boundary defined, we can explore the behaviors exhibited by these two categories of faults, as shown in Figure 2b, c (see Figure 3 for standardized CoCo values of all faults).

5 Comparison of geologic slip rates and geodetically based slip-deficit rates on strike-slip faults

Figure 2 illustrates the comparison between geologic and geodetic slip-deficit rates for the 24 different sites on the studied strike-slip faults. It reveals marked differences in the consistency of the values of the geodetic/geologic-rate pairs for high-CoCo faults relative to low-CoCo faults. Specifically, comparison of geodetic slip-deficit rates with large-displacement and small-displacement average geologic slip rates (displayed as squares and circles, respectively, in Figure 2) reveals that these rates are similar for faults characterized by low CoCo values (displayed in blue in Figure 2), whereas they differ for the faults characterized by high CoCo values (displayed in red in Figure 2). This observation is a corollary to the main conclusion of our previous study (Gauriau and Dolan, 2021), in which we showed that low-CoCo faults slip at relatively constant rates through time whereas high-CoCo faults exhibit long-term slip rates that are potentially different from the slip rates averaged over small displacements. In other words, the displacement over which the slip rate is averaged does not matter for low-CoCo faults, since any geologic slip rate will give the same value. In contrast, geologic slip rates for high-CoCo faults that are averaged over one particular displacement range may differ from the slip rate averaged over a different displacement range.

Figure 2a shows a comparison of geologic slip rates and geodetic slip-deficit rates. Figure 2b shows that low-CoCo strike-slip fault sites plot on (or near) the 1:1 line, reflecting the similarity of their short-term geodetic strain accumulation rates and both their small-displacement and large-displacement geologic strain-release rates. This can be further illustrated statistically, since the coefficient of determination obtained from an ordinary least squares regression for the low-CoCo faults is 0.983 for geologic rates averaged over large displacements, and 0.978 for geologic rates averaged over small displacements (Figure S1). Assuming a linear relationship between geologic slip rates and geodetic slip-deficit rates going through the origin, we find scaling lines with best-fit slopes and respective 1σ confidence of 0.945 ± 0.028 and 1.103 ± 0.050 for the low-CoCo faults using the large-displacement and small-displacement average geologic rates, respectively (see Figure S1a and b). These results show that for these low-CoCo faults, geodetic rates provide a reliable proxy for the geologic slip rate of the fault of interest.

That geodetic slip-deficit rates are a reliable proxy for geologic slip rate is not the case for high-CoCo faults (Figure 2c). Specifically, there is wide dispersion amongst the geodetic slip-deficit and both large- and small-displacement geologic slip rates (Figure 3). This observation requires that geodetic slip-deficit rates cannot be used as a proxy for geologic rates for high-CoCo faults, whether the rate is averaged over small displacements or large displacements. For these high-CoCo faults, the coefficient of determination obtained from an ordinary least squares regression between geologic

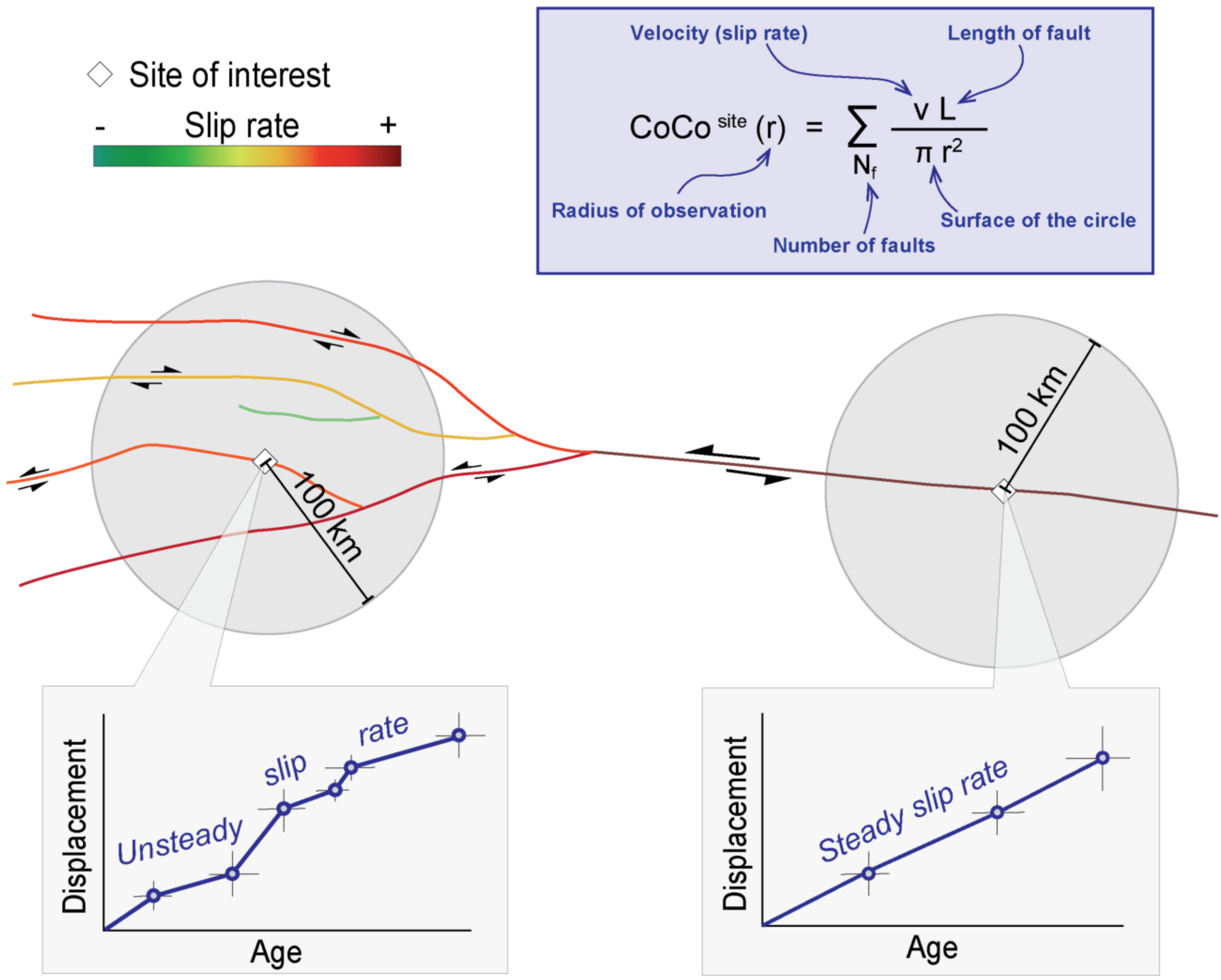


Figure 1 Schematic explanation of the rationale of the Coefficient of Complexity (CoCo) analysis for a hypothetical fault network. The calculation of CoCo for a given radius is shown on top. The radius over which CoCo is calculated is 100 km. Within a structurally complex fault system (numerous, and relatively fast-slipping faults), shown to the left, the CoCo value will be higher than within a structurally simple fault system (few or zero neighboring faults), shown to the right. The quantification of complexity, done with the CoCo analysis, correlates with the relative steadiness of geologic slip-rate record, as shown in our recent study (Gauriau and Dolan, 2021).

rates and geodetic slip-deficit rates is 0.396 for geologic rates averaged over large displacements, and 0.350 for geologic rates averaged over small displacements (Figure S1c, d). Scaling lines between geologic rates and geodetic rates for these faults, assuming a linear relationship going through the origin (as in Meade et al., 2013) are characterized by the best-fit slopes of 0.751 ± 0.162, using the small-displacement geologic rates, and 0.696 ± 0.140, using the large-displacement geologic rates (Figure S1c and d). These linear regressions seem to imply a global trend where geologic slip rates are faster than geodetic slip-deficit rates, but we suggest that these best-fit slope values are not meaningful, and are rather artifacts of the current limited state of available data. Reinforcing this idea is the observation that the dispersion of the data, shown by the standard deviations of the best-fit slopes, demonstrates that there is no good correlation between geodetic slip-deficit and geologic slip rates for high-CoCo faults. Figure 3 further illustrates this result, by displaying the ratio between the geodetic slip-deficit rates and the geologic rates averaged over large or small displacements. Figure 3b plots a measure of distance from the data points to the 1:1 ratio line with varying CoCo values, and emphasizes the dispersion of the data for higher-CoCo faults (see details of the dispersion calculation in the Supplementary Materials); the relatively sharp increase in dispersion at standardized CoCo ~0.0015-0.002 yr^{-1} likely reflects the presence of major secondary faults that can accommodate significant portion of relative plate motions.

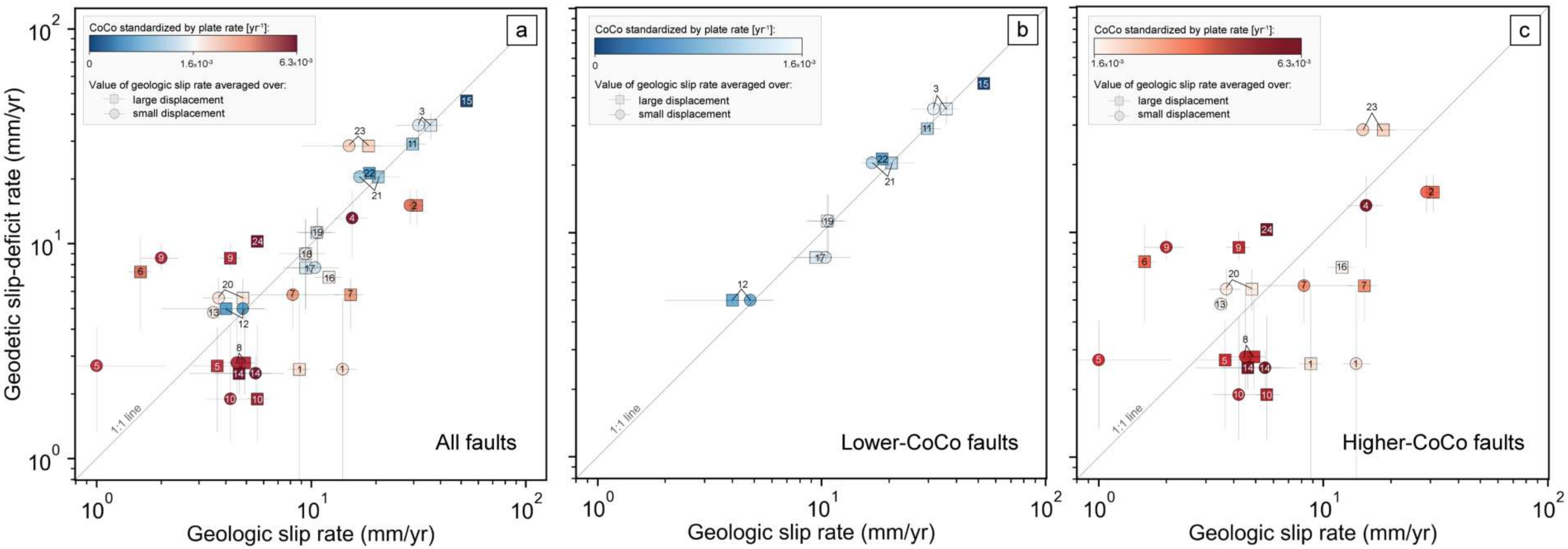


Figure 2 Geodetic slip-deficit rate and geologic slip-rate comparisons for major strike-slip faults. The geologic rates are shown as either averaged over a large displacement, or over a small displacement. The data points are color-coded according to their respective values of the Coefficient of Complexity (CoCo), standardized by the plate rate contained within a 100 km radius, as defined in Gauriau and Dolan (2021). The strike-slip faults considered in this study are: (1) Garlock, (2) San Andreas, Mojave segment, (3) San Andreas, Carrizo Plain segment, (4) San Jacinto, Claremont segment, (5) Owens Valley, (6) Calico, (7) Hope, (8) Wairau, (9) Clarence, (10) Awatere, (11) Alpine, (12) Dead Sea, Wadi Araba Valley, (13) Dead Sea, Beteiha, (14) Yammouneh, (15) Queen Charlotte, (16) Denali, central section, (17) Denali, western section, (18) Altyn Tagh, (19) Kunlun, (20) Haiyuan, (21) North Anatolian, Demir Tepe, (22) North Anatolian, Tahtaköprü, (23) Northern North Anatolian, (24) East Anatolian, Pazarcık (references listed in Table 1). **(a)** shows all the compiled faults in the same diagram. **(b)** shows all faults characterized by CoCo values that are less than 0.0016 yr^{-1} (referred to as low-CoCo faults). **(c)** shows all faults characterized by CoCo values that are more than 0.0016 yr^{-1} (referred to as high-CoCo faults).

6 Fault loading rates...

6.1 ...are constant on low-CoCo faults

Our analysis reveals that low-CoCo faults are characterized by geodetic rate/geologic rate ratios very close to one, regardless of the displacement scale over which the geologic slip rate is measured (Figures 2, 3). Geologic slip rates estimated from offset landforms at widely different displacements are the same for these faults, showing that the elastic strain release remains constant over the time intervals over which these displacements have accumulated. Furthermore, the current elastic strain accumulation rate (as constrained by the geodetic slip-deficit rate) is equal to strain release rates (as constrained by geologic slip rates) at all measured displacement scales. This indicates that for these faults, the elastic strain accumulation rate provided by the geodetic slip-deficit rate remains constant during the interseismic period (Figure 4), following the short-duration periods of fast post-seismic deformation at the beginning of each cycle, as originally noted by Meade et al. (2013).

6.2 ...vary on high-CoCo faults

In contrast, high-CoCo faults, embedded within more complex structural settings, display no consistent relationship between geodetic slip-deficit and geologic slip rates. As noted above, these results reinforce the point that geodetic slip-deficit rates cannot be used as reliable proxies for geologic slip rates on high-CoCo faults. Moreover, although the mismatch between geodetic slip-deficit rates and small-displacement geologic slip rates could conceivably be due to short-term variations in fault slip rate, the mismatch between geodetic slip-deficit rates and large-displacement geologic slip rates, which are averaged over >50 to hundreds of meters of slip (see Table 1) and numerous individual earthquakes, and will thus average over any shorter-term/smaller-displacement accelerations or decelerations of fault slip, indicates that elastic strain accumulation rates on the high-CoCo faults must vary through time. Specifically, at these large-displacement scales, the fault slip rate spanning numerous earthquakes will provide a robust estimate of the average rate of strain release on that fault through time. Insofar as the elastic strain accumulation rate must equal the elastic strain release rate (i.e., fault slip) over long time intervals, the mismatch that we document between geodetic slip-deficit rates and geologic slip rates averaged over large displacements requires that elastic strain accumulation rates as measured by geodetic slip-deficit rates must vary through time.

Further examination of the results displayed in Figure 3 helps us distinguish several types of behaviors amongst the high-CoCo faults. Those behaviors can be defined depending on whether the geodetic slip-deficit rate is equal to, slower than, or faster than either the large-displacement average geologic rate, or the small-displacement average geologic rate (Figure 4).

These differences between geodetic and geologic rates reveal the following fundamental point: Faults for which the current loading rate does not equal the average large-displacement geologic slip rate overly a ductile shear zone that must be creeping at either a slower or a faster rate than the long-term average slip rate. If, furthermore, the geodetic rate differs from the small-displacement rate, the rate of elastic strain accumulation consequently has to vary over the same periods of accelerations and decelerations that are averaged over in these small-displacement geologic rate values.

We suggest that using the mismatches between geodetic slip-deficit and small-displacement geologic rates can help us infer the current behavior of the faults that may be most representative of the near-future likelihood of major earthquake recurrence. Mismatches between elastic strain accumulation rates and small-displacement geological rates reveal three different modes for the high-CoCo faults. These are: faults that are storing elastic strain energy more slowly than their small-displacement geologic slip rates; faults that exhibit a current rate of elastic strain accumulation that is faster than the small-displacement geologic slip rate; and faults in which the geodetic slip-deficit rate approximately equals the youngest average geologic slip rate. In the following, we describe the details of the behavior of faults that fall within these three categories and discuss a model that attempts to explain the observations in terms of faults switching from one mode to another.

In the first case, geodetic slip-deficit rates are slower than the small-displacement (short-term) geologic slip rates measured on these faults. The Garlock (numbered 1 in Figures 2 and 3), the Mojave segment of the San Andreas (2), Wairau (8), Hope (9), Awatere (10), and Yammouneh (14) faults are all characterized by geodetic rate values that are slower than their respective geologic slip rates (both large- and small-displacement). For example, the central Garlock fault experienced a cluster of four large earthquakes between 0.5 and 2.0 ka (Dawson et al., 2003), resulting in a small-displacement (26 m) slip rate averaged over these four events through to the present of $14^{+2.2}_{-1.8}$ mm/yr (Dolan et al., 2016). Modeling of geodetic data consistently yields very slow rates of elastic strain accumulation on the central Garlock fault, with a best estimate of ~2.6 mm/yr (Evans, 2017b), potentially including almost no elastic strain accumulation. In contrast, the large-displacement (long-term) slip rate averaged over the most recent 70 m of slip on this section of the Garlock fault is 8.8±1.0 mm/yr (Fougere et al., 2023, submitted). While this is slower than the small-displacement geologic rate, it is at least three times faster than the current rate of elastic strain accumulation. This mismatch suggests that the Garlock fault has recently entered into a “slow” mode of elastic strain accumulation, likely as a result of a decreased shearing rate on the underlying ductile shear zone. But why is the youngest, small-displacement rate so fast? We suggest that the switch in behavior of the Garlock fault from the 0.5 – 2 ka “fast” mode ended with the final earthquake in the cluster, either because the fault (including the upper seismogenic part and the ductile shear zone roots) strengthened during the fast period encompassing the four-event cluster and became more difficult to slip (Dolan et al., 2007; Cawood and Dolan,

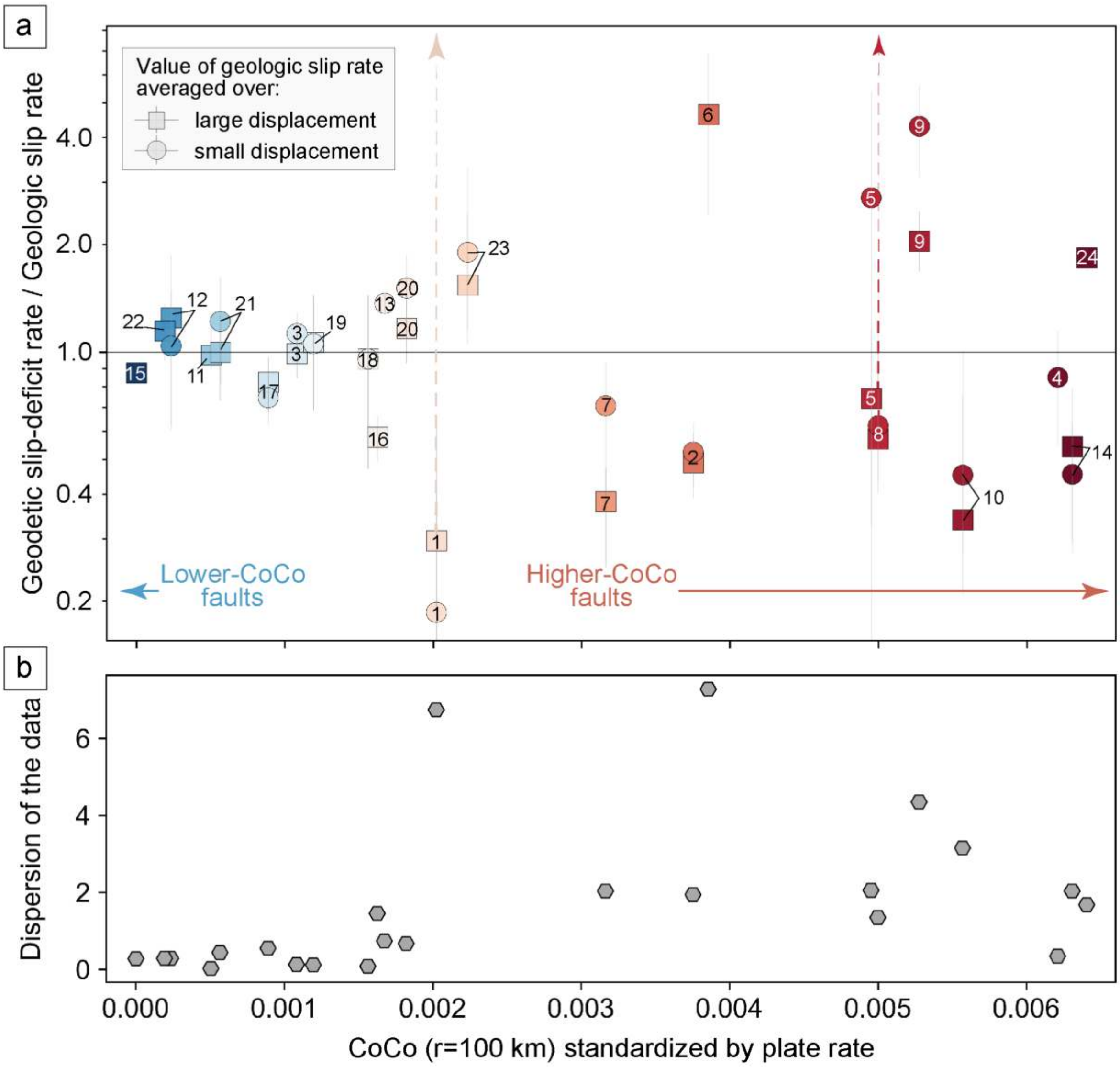


Figure 3 Variations of geodetic to geologic slip-rate ratios against CoCo values standardized by plate rate over a 100 km radius. (**a**) Ratios of geodetic slip-deficit rate to geologic rate plotted against CoCo. The geologic rate values are averaged over large or small displacement (as in Figure 2). The numbering of the fault sites is referred to in Figure 2 and Table 1. The dashed arrows refer to a ratio of geodetic/geologic rate that would reach infinity, with a geological rate close or equal to 0 mm/yr, if the fault has not slipped for a long time since the MRE (see text for details). (**b**) Diagram showing the dispersion of the ratio (geodetic to geologic rates) values varying with the CoCo values. The higher the CoCo value, the more scattered the data (i.e., the farther from the 1:1 ratio line they tend to plot). The measure of the dispersion is detailed in the Supplementary Materials. Although we cannot calculate an exact CoCo value for the Queen Charlotte fault (15), because of our inability to include all active faults within a 100 km radius of the slip-rate site, we assign it a CoCo value of zero, since this fault accommodates >95 % of the total Pacific/North America plate-motion rate (NUVEL-1A; DeMets and Dixon, 1999).

submitted), and/or because the Garlock fault has exhausted what Dolan et al. (2024) refer to as the "crustal strain capacitor" (similar to Mencin et al. (2016) "strain reservoir"), that is, the shear strain stored in the crust surrounding this section of the Garlock fault. In this view, the current slip rate (or, equivalently in this context, the "most recent geologic slip rate") of the Garlock fault since the most recent earthquake (MRE) ca. 500 years ago has been 0 mm/yr, reflecting the current very slow rate of elastic strain accumulation on the Garlock fault.

Similarly, the geodetic slip-deficit rate on the Wairau fault in New Zealand ($2.8^{+2.4}_{-0.8}$ mm/yr; Johnson et al., 2022) is slower than the small-displacement rate of 4.5±1.0 mm/yr (Zinke et al., 2021), calculated for the preceding fast period of slip between a geomorphic offset dated at ca. 5.4 ka and the ca. 2 ka MRE. This contrast highlights a period of fast slip on the fault during this time interval. Yet, 2,000 years have elapsed since the MRE on the Wairau fault (relative to an average Holocene recurrence interval of ca. 1,000 years Nicol and Dissen, 2018), which we suggest indicates a "most recent geologic slip rate" since the MRE of 0 mm/yr. Thus, the averaging of the small-displacement rate over

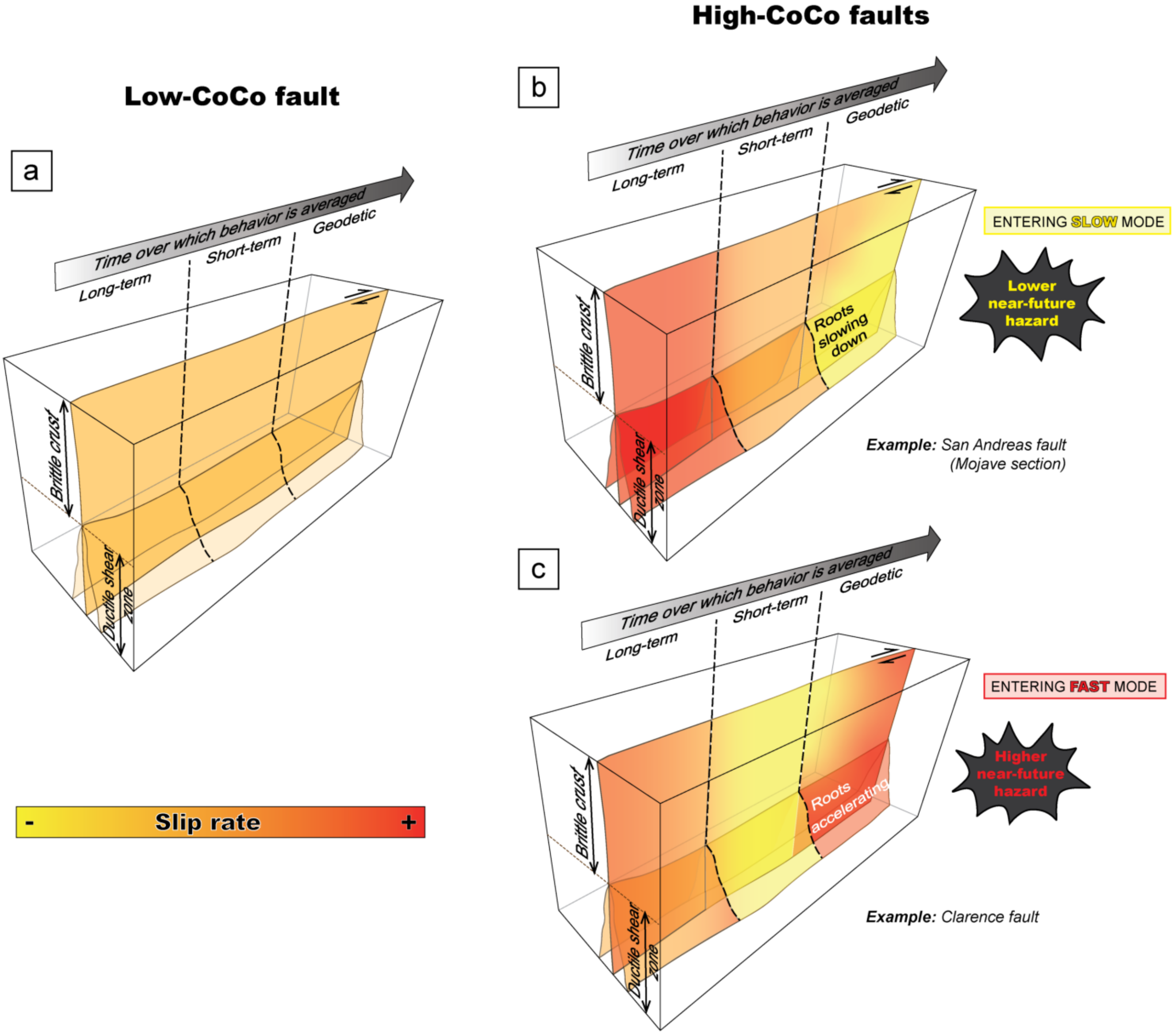


Figure 4 Observed modes of fault behavior, with time shown as the horizontal dimension of the block, and with relative slip rate displayed with a color gradient. In **(a)**, we show that whatever the time over which its behavior is averaged, a low-CoCo fault's slip rate is constant and thus equals its elastic strain accumulation rate, as shown in the left hand-side, hence the same color at each point in time and in the brittle and ductile parts of the fault. Note that we are not considering single-earthquake time scales. In contrast, high-CoCo faults (**b** and **c**) exhibit several types of behaviors, as discussed in the text. In **(b)**, we illustrate a fault that has a short-term (small-displacement) geologic slip rate that is slower than its long-term (large-displacement) rate. For this fault, the current elastic strain accumulation (ductile shear of the ductile roots) is slower than the short-term geologic slip rate, and therefore might be entering what we refer to as a slow mode. In **(c)**, we show another example of a fault whose long-term geologic slip rate is faster than its short-term geologic slip rate. This fault is entering a fast mode since its elastic strain accumulation is much faster than its short-term geologic slip rate.

the past 5,400 years through to the present may be masking a switch of the Wairau fault from a fast mode between 2 and 5 ka, to the current slow mode that has prevailed since the MRE at 2 ka. In both the Wairau and Garlock faults examples, if we were to use the inferred most recent geologic slip rate of 0 mm/yr as the best representation of the small-displacement slip rate, the geodetic/small-displacement rate ratios would soar, as the dashed arrows in Figure 3a illustrate.

Another example is the Yammouneh fault (14), which has a geodetic slip-deficit rate (2.5±0.5 mm/yr; Gomez et al., 2020) that is much slower than its small-displacement slip rate (5.5±2.0 mm/yr; Daëron et al., 2004) (Figure 3). The Yammouneh fault might therefore also be experiencing a slow mode since the MRE in 1202 C.E. (Daëron et al., 2007, Table S1).

Although the small-displacement slip-rate of the Hope fault ($8.2^{+5.4}_{-3.0}$ mm/yr; Hatem et al., 2020) is likely faster than the geodetic slip-deficit rate estimate ($5.8^{+1.8}_{-1.1}$ mm/yr; Johnson et al., 2022), their respective 2σ uncertainties overlap (Table 1), which does not allow us to strongly affirm a potential switch of mode for this fault. However, the difference between these estimates might suggest that the Hope fault is currently in a slower

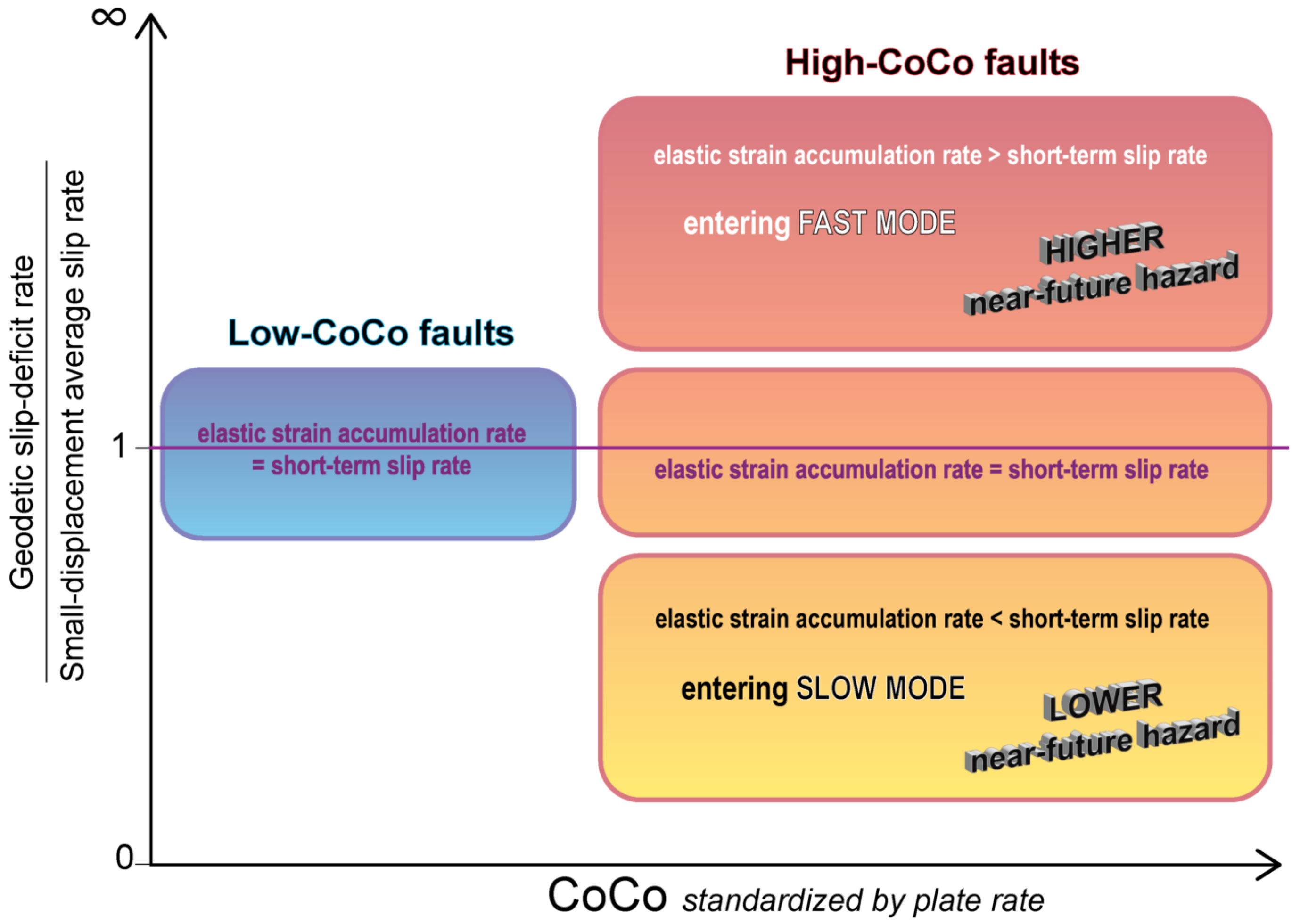


Figure 5 Schematic illustration of modes of behavior defined in this paper, according to the CoCo values and the geodetic/geologic rate ratio, and their potential meaning in terms of near-future hazard.

mode, and may have exhausted its strain capacitor in the past five earthquakes, which generated 20-30 m of fault slip over the past ~1,500 years (Hatem et al., 2019, 2020). The thus-reduced shear stress stored in the crust surrounding the Hope fault might explain the lack of significant slip on the Hope fault in the 2016 Kaikōura earthquake sequence (e.g., Hamling et al., 2017), despite its proximity to the faults that initially ruptured in the sequence. Indeed, both Ulrich et al. (2019) and Nicol et al. (2023) have suggested that the lack of significant 2016 coseismic slip on the Hope fault could be due to the low stresses in play across the Hope fault prior to the Kaikōura earthquake.

A final example is the Mojave section of the San Andreas fault (SAFm), which is characterized by an elastic strain accumulation rate (15.1±2.3 mm/yr; Evans, 2017b) that is much slower than its small-displacement slip rate (~27-29 mm/yr; Weldon et al., 2004; Dolan et al., 2016) (Figures 3, 4, Table 1). The MRE occurred 167 years ago on the SAFm, whereas the mean recurrence interval for this stretch of the fault is about 100 years (e.g., Scharer et al., 2017). The absence of any earthquakes since the 1857 MRE led to much speculation in earlier decades, when some scientists suggested that the SAFm was "overdue" (e.g., Weldon and Sieh, 1985). These early ideas of earthquake recurrence patterns were based on the assumption of steady elastic strain accumulation rates. If, instead, elastic strain accumulation rates vary, as we show here, then the long elapsed time since the 1857 earthquake may at least partially be a consequence of reduced loading rates in this section of the SAF, as reflected in the current geodetic rate. All of this suggests that the SAFm (2) may have entered a "quieter mode".

A partial, potential alternative explanation for this situation was provided in Hearn et al. (2013) and Hearn (2022), who suggested that some of this slow elastic strain deformation rate on the SAFm might be due to a so-called "ghost transient" related to long-term viscoelastic relaxation of the lithospheric mantle and lower crust following the 1857 Fort Tejon earthquake. However, this would only explain 5 mm/yr of the apparent ~14 mm/yr difference between the geodetic slip-deficit rate and the small-displacement slip rate. In marked contrast to the SAFm, Hearn et al. (2013) also noted that there is no such "ghost transient" associated with the Garlock fault, which ruptured most recently in 1450-1640 CE (Dawson et al., 2003).

Our analysis reveals another type of behavior, in which faults exhibit geodetic slip-deficit rates that are faster than their geologic slip rates. We suggest that these faults may have switched from a slow mode to a fast mode. This behavior characterizes the Clarence fault (9), the northern Dead Sea fault (nDSF - 13), the northern strand of the North Anatolian fault system (nNAF - 23), and the Pazarcık segment of the East Anatolian fault (EAF – 24) (Figure 3). The Clarence fault (9) has a geodetic slip-deficit rate ($8.6^{+1.5}_{-1.1}$ mm/yr; Johnson et al., 2022) that is faster than both its small-displacement and large-displacement geologic rates, although its small-displacement slip rate (2.0 ± 0.4 mm/yr) is half as fast as its large-displacement slip rate (4.2 ± 0.5 mm/yr; Zinke et al., 2019). Similarly, the nDSF stores elastic strain energy at a rate of 4.8 ± 0.3 mm/yr (Gomez et al., 2020) and is characterized by a slower small-displacement slip rate of 3.5 ± 0.2 mm/yr (Wechsler et al., 2018). For the nNAF, considering the large uncertainties on the large-displacement geologic slip rate ($18.5^{+10.9}_{-5.9}$ mm/yr, measured over a 500 My time scale; Kurt et al., 2013), we cannot confidently infer that it is slower than the reported geodetic slip-deficit rate (28.6 mm/yr; DeVries et al., 2016), but we can more confidently state that the small-displacement geologic rate (15 ± 6 mm/yr; Meghraoui et al., 2012) is slower that the geodetic rate, as suggested by Dolan and Meade (2017). The EAF (24) has a geodetic slip-deficit rate (10.3 ± 0.6 mm/yr; Aktug et al., 2016) that is nearly twice as fast as the available large-displacement geologic slip rate (5.6 ± 0.3 mm/yr; Yönlü and Karabacak, 2023). Notably, this section of the EAF ruptured in the 2023 M_w 7.8 Kahramanmaraş earthquake.

The Calico fault (6) may also fall within this type of behavior, with a switch from a previous slow mode to a current faster mode. Although the data currently available for the Calico fault do not allow us to infer a small-displacement slip rate, the current loading rate (7.4±3.4 mm/yr; Evans, 2017b) is much faster than its large-displacement slip rate (1.6±0.2 mm/yr; Oskin et al., 2007) (Figure 3). Specifically, the Calico fault has generated four surface-rupturing earthquakes within the past ~9,000 years (Ganev et al., 2010), which coincide with periods of clustered moment release identified on other faults in the eastern California shear zone (ECSZ) (Rockwell et al., 2000). The MRE on the Calico fault occurred sometime between 0.6 and 2 ka, likely as part of an ongoing cluster of earthquakes that has been occurring over the past 1-1.5 ky in the ECSZ (Rockwell et al., 2000), including most recently the 1872 Owens Valley, 1992 Landers, 1999 Hector Mine, and 2019 Ridgecrest earthquakes. Geodetic data suggest that the Calico fault, and potentially other nearby faults in the ECSZ, are likely experiencing a period of anomalously fast loading (Oskin et al., 2007; Dolan et al., 2007), as originally suggested by Peltzer et al. (2001), and further discussed by Oskin et al. (2008). Peltzer et al. (2001) showed that active dextral shear associated with the ECSZ extends across the Garlock fault, which does not exhibit any accumulation of left-lateral shear strain energy, emphasizing the idea that the Garlock fault has entered a slow mode (Evans et al., 2016; Evans, 2017a). These observations are consistent with kinematic models that suggest that the Garlock fault is currently storing and releasing elastic strain energy at much slower-than-average rates, whereas the ECSZ subsystem is storing and releasing energy at faster-than-average rates (Dolan et al., 2007, 2016; Hatem and Dolan, 2018; Peltzer et al., 2001). Farther north in the ECSZ-Walker Lane system, the Owens Valley fault exhibits a geodetic slip-deficit rate estimate (2.7±1.4 mm/yr; Evans, 2017b) that may be faster than its small-displacement slip-rate (1.3±0.8 mm/yr; Haddon et al., 2016), consistent with a period of faster-than-average elastic strain accumulation. It is worth noting however, that these rate estimates overlap at 95% uncertainty (Table 1).

In addition to these behaviors, the San Jacinto fault (4) exhibits a small-displacement geologic slip rate (15.6±2.3 mm/yr; Onderdonk et al., 2015) that is similar to the current loading rate (13.2±4.6 mm/yr; Evans, 2017b) within 2σ uncertainties. However, there is currently no well-constrained, large-displacement (> 50 m) geologic slip rate available for the San Jacinto fault. Thus, the similarity of the geodetic and small-displacement geologic rates might suggest that the San Jacinto fault may have been captured in the middle of either a fast period (i.e., cluster) or a slow period, but in the absence of a large-displacement slip rate, we cannot say definitively which.

It is worth noting that the slip rate of high-CoCo faults does not seem to affect their behavior; both fast-slipping and slow-slipping high-CoCo faults exhibit significant dispersion of geodetic/geologic ratios. Dispersion analysis indicates that fast-slipping, high-CoCo faults exhibit larger dispersion of geodetic/geologic ratios than for slower-slipping high-CoCo faults (see Supplementary Materials), contrary to what Cowie et al. (2012) obtained from their simulations of elastic interactions between growing faults. However, we suspect that the dispersion values we determine are not particularly meaningful given the dearth of slip-rate data from fast-slipping, high-CoCo faults.

One key element to highlight is the potential difficulty in capturing any switches from fast to slow mode (or vice versa) with the available incremental fault slip-rate data, which in some instances may not be detailed enough over the appropriate displacement intervals to capture these switches in mode. This challenge will typically lie in the resolution at which the increments of the incremental slip-rate record are obtained, and if the slip-rate data are not detailed enough over the appropriate time and displacement intervals, the switches in mode may not be observable. Assuming, however, that the input data we use in this study provide sufficient information to constrain the timing of these switches in mode, our results imply that the elastic strain accumulation rate keeps up with or controls fast and slow fault slip periods, which challenges the suggestion by Weldon et al. (2004) that the strain release rate varies while the strain accumulation rate does not (i.e., their "strain-predictable behavior").

7 Ductile shear zone behavior...

7.1 ...on high-CoCo faults

The variations in strain accumulation rate described above likely record variations in the rate of shear along the ductile shear zone roots of seismogenic faults. Here we discuss the mechanisms that might control the behavior of ductile shear zones on high-CoCo faults.

The different behaviors exhibited by the high-CoCo faults can be explained by mechanisms that occur at the plate-boundary scale, such as the shared accommodation of slip in complex plate-boundary structural settings (Peltzer et al., 2001; Dolan et al., 2016), as well as by mechanisms at the scale of the fault zone, with potential strengthening and weakening processes over the ductile shear zone and the coupling between the brittle and the ductile parts of a fault (e.g., Peltzer et al., 2001; Oskin et al., 2008; Dolan et al., 2007). In structurally complex, high-CoCo settings, mechanically complementary faults within the system can share the load by trading off slip while maintaining a relatively constant overall system-level rate that keeps pace with the relative plate-motion rate (Dolan et al., 2024). In these structurally complex plate-boundary fault systems, when one fault slips much faster than its average rate throughout multiple earthquakes, the other faults of the system slip more slowly or not at all as the overall fault system works together to maintain constant average rate. Acceleration of the ductile shear zone rate will create a positive feedback loop in which faster shear on the ductile shear zone roots will drive the occurrence of more frequent, large earthquakes (i.e., an earthquake cluster) in the seismogenic part of the fault, which will in turn accelerate underlying ductile shear rates through viscous coupling, increasing driving stress, and potentially by addition of fluids into the nominally ductile uppermost parts of the ductile shear zone roots (Ellis and Stöckhert, 2004; Cowie et al., 2012; Mildon et al., 2022; Dolan et al., 2007). But eventually, either through exhaustion of the crustal strain capacitor of stored elastic strain energy on the fault in question, and/or through increases in ductile shear zone strength (i.e., resistance to shear), the fault will enter a slow mode of strain release as deformation shifts to a mechanically complementary, weaker fault within the system (Dolan et al., 2024).

These accelerations and/or decelerations of the faults' ductile shear roots of a complex fault network might be explained by strength changes (e.g., strain hardening and weakening). Dolan et al. (2007, 2016) and Dolan and Meade (2017), for instance, suggested that ductile shear zone roots can harden during fast slip periods, leading to lulls in ductile shear and hence earthquake lulls in the upper crust. In this model, the ductile shear roots of faults are accumulating elastic strain energy more slowly than their long-term slip rate, after having been "exhausted" during a period of rapid ductile shearing and fast fault slip in clusters of earthquakes (Dolan et al., 2024). Other potential mechanisms occurring within ductile shear zones that could give rise to a change in shearing rate and associated elastic strain accumulation rates of the overlying fault include changes in fluid concentration (e.g., Mancktelow and Pennacchioni, 2004; Okazaki et al., 2021), changes in grain size (e.g., Handy, 1989; Okudaira et al., 2017), macroscopic fault evolution (e.g., Handy et al., 2007) and fabric development (e.g., Carreras et al., 2005; Melosh et al., 2018) (see Cawood and Dolan, submitted, for details on these mechanisms). All these mechanisms could drive the crustal "strain capacitor" to either its exhaustion or its replenishment (Dolan et al., 2024; Cawood and Dolan, submitted).

7.2 ...on low-CoCo faults

In contrast, tectonically isolated, primary low-CoCo plate-boundary faults (e.g. central SAF, central and eastern NAF, Alpine fault), are characterized by interseismic rates that correlate well with geologic slip rates that are averaged over both small and large displacements (Figure 3). This suggests that such low-CoCo faults must "keep up" with the relative plate-motion rate over short time and small displacement scales because there are no other mechanically complementary faults in such systems to share the load. In other words, even though all of the potential strengthening and weakening mechanisms we discuss for high-CoCo faults must be operating on low-CoCo faults as well, these processes will be overwhelmed by steady increases in driving stress related to relative plate motion. All or most of the relative plate motion must be accommodated on the primary fault in the absence of other major faults that could potentially share the work required to move the plates past each other. Moreover, the similarity of geodetic slip-deficit rates and small-displacement geologic slip rates on low-CoCo faults requires that the fault responds to steady increases in driving stress at scales of no more than a few tens of meters of relative plate motion. This is consistent with the long-held notion embodied in elastic rebound theory (Reid, 1910) that the crust can only store a given amount of elastic strain energy before the weakest element of the system (i.e., the structurally isolated primary fault) slips in an earthquake. In turn, this line of reasoning implies that the single, isolated fault either has to be weak all the time - as soon as it stores no more than a few tens of meters of elastic strain energy, it is ready to slip - or it cyclically becomes weak when stress is approaching the rupture limit. A key question is whether this near-1:1 relationship between "energy in" (as manifest in geodetic slip-deficit rates) and "energy out" (i.e., fault slip rates) on low-CoCo faults extends to single-earthquake scales. The few available earthquake-by-earthquake age plus displacement-per-event datasets that are available from low-CoCo faults suggest that, at least generally, this may be the case. Specifically, the relatively regular timing (CoV ~ 0.3) of surface ruptures on the Alpine fault at Hokuri Creek, coupled with similar ~7.5 m horizontal displacements in the two most recent earthquakes (Berryman et al., 2012; De Pascale and Langridge, 2012; Sutherland et al., 2006), and the similar displacements in the four most recent earthquakes and relatively regular timing of earthquakes on the NAF at Demir Tepe (Kondo et al., 2010) are consistent with the idea that this may extend to single earthquake scales. If this is generally true, then low-CoCo faults may release much of, and perhaps almost all, of the shear stress accumulated since the previous event during each rupture. It is worth noting, however, that even at the Hokuri Creek site on the low-CoCo Alpine fault (Berryman et al., 2012), which is characterized by quasi-periodic earthquake recurrence, the 24-event record cannot be fit precisely with either time- or slip-predictable models (Shimazaki and Nakata, 1980), and may best be explained by an underlying chaotic behavior (Gauriau et al., 2023).

8 Fault's near-future behavior, and further applications for PSHA

Our results may provide new insight into how slip rates can be better used as basic inputs into probabilistic seismic hazard assessment (PSHA) methods. For low-CoCo faults, the outcome is straightforward – both the slip rate averaged over large displacements and the slip rate averaged over small displacements are similar to the geodetic slip-deficit rate. Therefore, any of these values can be used as an input into a PSHA. Despite this relative constancy of both strain accumulation and release rates in the behavior of a low-CoCo fault, any attempt towards formulating earthquake prediction focused on timing of earthquake occurrence on a specific fault may be functionally impossible (e.g., Chen et al., 2020; Gauriau et al., 2023). Therefore, a probabilistic methodology is required for any seismic hazard assessment.

For high-CoCo faults, the outcome is less straightforward, since such faults exhibit variable strain accumulation and release rates through time. The question arises as to what slip-rate value is the best to use in PSHA? There are three possible strategies for incorporating incremental slip-rate data into PSHA, as originally suggested by van Dissen (2020): (a) incorporating the large-displacement average slip rate by neglecting any incremental rate changes, which in a long-term statistical sense can be viewed as variations about the mean rate; (b) using the full error range associated with all available incremental slip rates, or (c) favoring the most recent (smallest-displacement multiple-earthquake) incremental slip rate as the most appropriate one.

Here we propose a potential solution to this conundrum by comparing the small-displacement and large-displacement rates with the elastic strain accumulation rates. Geodetic slip-deficit rates have been suggested as primary inputs into seismic hazard assessment (e.g., Bird and Kreemer, 2014; Hussain et al., 2018), but never in light of comparison to available geologic slip-rate records. The examples listed in paragraph 6.2., however, illustrate the current limitations on using small-displacement rates (suggestion c) as a proxy for the most recent phase of fault behavior without considering the possibility that the fault may have switched modes in the interval since displacement of the most-recent available small-displacement slip-rate data. We suggest that a potential path forward is to use the comparison of the geodetic slip-deficit rates with small-displacement geologic rates of high-CoCo faults to forecast the near-future behavior that might be expected on a given fault. While we suggested in our earlier paper (Gauriau and Dolan, 2021) that option (c), i.e., implementing the shorter-term slip rate into a PSHA, would lead to a more reliable forecast of the near-future behavior of the fault, the current analysis suggests that deviations of geodetic rates from the small-displacement geologic slip rates might better illustrate the future behavior of high-CoCo faults.

Specifically, we propose that a geodetic slip-deficit rate that is slower than the small-displacement slip rate might indicate lower near-future hazard, because the fault is storing elastic strain energy more slowly than average (Figure 5). This is exemplified by the cases of the Garlock fault, the SAFm, and the Hope fault. Conversely, geodetic rates that are faster than the small-displacement rate on faults that have not experienced a recent earthquake (i.e., those not experiencing a post-seismic strain transient) may indicate higher near-future hazard, as illustrated by the nNAF, the Clarence fault, and the nDSF. In support of this idea, the 2023 M_w 7.8 Kahramanmaraş earthquake occurred on a section of the EAF that exhibited a geodetic slip-deficit rate, prior to the earthquake, that was almost twice as fast as the long-term geologic slip rate. In the case of the San Jacinto fault, and other faults with a geodetic rate that equals the small-displacement slip rate, we suggest that the near-future hazard can be best represented by the small-displacement slip rate and/or the geodetic rate (Figure 5).

One possible route towards using these observations in improved PSHA would be to evaluate geodetic and geologic rate discrepancies using the smallest-displacement incremental slip rate for a fault to infer the current mode of fault behavior.

9 Conclusions

Our comparison of geologic fault slip rates with geodetic slip-deficit rates from strike-slip plate-boundary faults reveals markedly different strain accumulation and release behavior on structurally isolated faults relative to those that extend through structurally complex regions. Our main take-away is that elastic strain accumulation rates on high-CoCo faults must vary through time, whereas they remain relatively constant on low-CoCo faults. This can potentially be applied to faults exhibiting other kinematics, such as extensional or compressional fault systems, where both fault interactions and slip-rate variability have also been studied (e.g., Luo and Liu, 2010; Mildon et al., 2022).

High-CoCo faults have geodetic-to-geologic ratios that vary widely, demonstrating that rates of elastic strain accumulation vary significantly through time at scales that are longer than individual earthquake cycles. This is particularly clear from the differences observed between the short-term geodetic slip-deficit rate data with long-term, large-displacement geologic slip rates, which will average over any shorter-term and smaller-displacement accelerations and decelerations of fault slip that typify faults in such settings (Gauriau and Dolan, 2021). Presumably, these changes reflect temporally variable rates of shear on the ductile shear zone roots of brittle faults, which we infer are related to the more complicated history of strain accumulation and release among regional fault interactions at displacement scales of a few tens of meters and centennial to millennial time scales. Specifically, geodetic slip-deficit rates that neither match large-displacement nor small-displacement average slip rates indicate that the elastic strain accumulation rate must vary over time scales corresponding to the deceleration and acceleration periods over which smallest-displacement geologic rates are averaged.

In contrast, low-CoCo faults are characterized by steady elastic strain accumulation and release rates, which indicate that such faults need to "keep up with" the relative plate motion rate at short-time and small-displacement scales, overwhelming any potential strengthening and weakening mechanisms that might be operating on such faults. Consequently, the geodetic slip-deficit rate observed on a low-CoCo fault can be used as a proxy for its geologic rate, which itself can be assumed to be relatively constant.

Finally, we suggest that the discrepancies between short-term geologic slip rates and geodetic slip-deficit rates for high-CoCo faults might represent a switch of mode, revealing either an accelerating or a decelerating phase. A geodetic slip-deficit rate that is faster than the most recent geologic incremental slip rate would imply a potential higher near-future seismic hazard, whereas a geodetic rate that is slower than the smallest-displacement slip rate would signal a lower near-future seismic hazard. These discrepancies could be used to refine PSHA models, not only in strike-slip fault systems, as highlighted in this study, but potentially to any type of plate-boundary kinematics. The importance and current relative dearth of robust incremental slip rate records highlights the need to develop more such records from more faults around the world to enable better PSHA.

Acknowledgements

We thank editor Ake Fagereng and reviewers Zoë Mildon and Sam Wimpenny for their constructive comments that helped improve the quality of this manuscript.

10 Data and code availability

The data used in this study, and necessary to reproduce our results, are all part of published articles, referred to in Table 1 and throughout the manuscript.

11 Competing interests

The authors have no competing interests.

References

Aktug, B., Ozener, H., Dogru, A., Sabuncu, A., Turgut, B., Halicioglu, K., Yilmaz, O., and Havazli, E. Slip rates and seismic potential on the East Anatolian Fault System using an improved GPS velocity field. *Journal of Geodynamics*, 94–95:1–12, Mar. 2016. doi: 10.1016/j.jog.2016.01.001.

Barbot, S., Luo, H., Wang, T., Hamiel, Y., Piatibratova, O., Javed, M. T., Braitenberg, C., and Gurbuz, G. Slip distribution of the February 6, 2023 Mw 7.8 and Mw 7.6, Kahramanmaraş, Turkey earthquake sequence in the East Anatolian Fault Zone. *Seismica*, 2(3), Apr. 2023. doi: 10.26443/seismica.v2i3.502.

Barka, A. The North Anatolian fault zone. *Annales tectonicae*, suppl. VI:164–195, 1992.

Barka, A., Akyüz, H., Altunel, E., Sunal, G., Cakir, Z., Dikbas, Aand Yerli, B., Armijo, R., Meyer, B., de Chabalier, J., Rockwell, T., Dolan, J., Hartleb, R., Dawson, T., Christofferson, S., Tucker, A., Fumal, T., Langridge, R., Stenner, H., Lettis, W., Bachhuber, J., and Page, W. The Surface Rupture and Slip Distribution of the 17 August 1999 Izmit Earthquake (M 7.4), North Anatolian Fault. *Bulletin of the Seismological Society of America*, 92(1): 43–60, Feb. 2002. doi: 10.1785/0120000841.

Barth, N., Kulhanek, D., Beu, A., Murray-Wallace, C., Hayward, B., Mildenhall, D., and Lee, D. New c. 270 kyr strike-slip and uplift rates for the southern Alpine Fault and implications for the New Zealand plate boundary. *Journal of Structural Geology*, 64: 39–52, 2014. doi: 10.1016/j.jsg.2013.08.009.

Bender, A. M., Lease, R. O., Rittenour, T., and Jones, J. V. Rapid active thrust faulting at the northern Alaska Range front. *Geology*, 51(6):527–531, Mar. 2023. doi: 10.1130/G51049.1.

Bendick, R., Bilham, R., Freymueller, J., Larson, K., and Yin, G. Geodetic evidence for a low slip rate in the Altyn Tagh fault system. *Nature*, 404(6773):69–72, Mar. 2000. doi: 10.1038/35003555.

Benedetti, L., Finkel, R., Papanastassiou, D., King, G., Armijo, R., Ryerson, F., Farber, D., and Flerit, F. Post-glacial slip history of the Sparta fault (Greece) determined by 36 Cl cosmogenic dating: Evidence for non-periodic earthquakes. *Geophysical Research Letters*, 29(8), Apr. 2002. doi: 10.1029/2001gl014510.

Berryman, K. R., Cochran, U. A., Clark, K. J., Biasi, G. P., Langridge, R. M., and Villamor, P. Major Earthquakes Occur Regularly on an Isolated Plate Boundary Fault. *Science*, 336(6089):1690–1693, June 2012. doi: 10.1126/science.1218959.

Bird, P. and Kreemer, C. Revised Tectonic Forecast of Global Shallow Seismicity Based on Version 2.1 of the Global Strain Rate Map. *Bulletin of the Seismological Society of America*, 105(1): 152–166, Dec. 2014. doi: 10.1785/0120140129.

Brothers, D. S., Miller, N. C., Barrie, J. V., Haeussler, P. J., Greene, H. G., Andrews, B. D., Zielke, O., Watt, J., and Dartnell, P. Plate boundary localization, slip-rates and rupture segmentation of the Queen Charlotte Fault based on submarine tectonic geomorphology. *Earth and Planetary Science Letters*, 530:115882, Jan. 2020. doi: 10.1016/j.epsl.2019.115882.

Bull, J., Barnes, P., Lamarche, G., Sanderson, D., Cowie, P., Taylor, S., and Dix, J. High-resolution record of displacement accumulation on an active normal fault: implications for models of slip accumulation during repeated earthquakes. *Journal of Structural Geology*, 28(7):1146–1166, July 2006. doi: 10.1016/j.jsg.2006.03.006.

Carreras, J., Druguet, E., and Griera, A. Shear zone-related folds. *Journal of Structural Geology*, 27(7):1229–1251, July 2005. doi: 10.1016/j.jsg.2004.08.004.

Cawood, T. and Dolan, J. Fault strength variations over multiple seismic cycles: The role of fluids, cementation, strain hardening, and shear folding, submitted.

Chen, Y., Liu, M., and Luo, G. Complex Temporal Patterns of Large Earthquakes: Devil's Staircases. *Bulletin of the Seismological Society of America*, 110(3):1064–1076, Apr. 2020. doi:

10.1785/0120190148.

Cowgill, E. Impact of riser reconstructions on estimation of secular variation in rates of strike–slip faulting: Revisiting the Cherchen River site along the Altyn Tagh Fault, NW China. *Earth and Planetary Science Letters*, 254(3):239–255, Feb. 2007. doi: 10.1016/j.epsl.2006.09.015.

Cowgill, E., Gold, R. D., Xuanhua, C., Xiao-Feng, W., Arrowsmith, J. R., and Southon, J. Low Quaternary slip rate reconciles geodetic and geologic rates along the Altyn Tagh fault, northwestern Tibet. *Geology*, 37(7):647–650, July 2009. doi: 10.1130/G25623A.1.

Cowie, P. A., Roberts, G. P., Bull, J. M., and Visini, F. Relationships between fault geometry, slip rate variability and earthquake recurrence in extensional settings: Fault geometry control on earthquake rupture. *Geophysical Journal International*, 189(1): 143–160, Feb. 2012. doi: 10.1111/j.1365-246x.2012.05378.x.

Daëron, M., Benedetti, L., Tapponnier, P., Sursock, A., and Finkel, R. C. Constraints on the post ~25-ka slip rate of the Yammoûneh fault (Lebanon) using in situ cosmogenic 36Cl dating of offset limestone-clast fans. *Earth and Planetary Science Letters*, 227 (1–2):105–119, Oct. 2004. doi: 10.1016/j.epsl.2004.07.014.

Daëron, M., Klinger, Y., Tapponnier, P., Elias, A., Jacques, E., and Sursock, A. 12,000-Year-Long Record of 10 to 13 Paleoearthquakes on the Yammouneh Fault, Levant Fault System, Lebanon. *Bulletin of the Seismological Society of America*, 97(3): 749–771, June 2007. doi: 10.1785/0120060106.

Daout, S., Jolivet, R., Lasserre, C., Doin, M.-P., Barbot, S., Tapponnier, P., Peltzer, G., Socquet, A., and Sun, J. Along-strike variations of the partitioning of convergence across the Haiyuan fault system detected by InSAR. *Geophysical Journal International*, 205(1):536–547, Apr. 2016. doi: 10.1093/gji/ggw028.

Dawson, T. E., McGill, S. F., and Rockwell, T. K. Irregular recurrence of paleoearthquakes along the central Garlock fault near El Paso Peaks, California. *Journal of Geophysical Research: Solid Earth*, 108(B7), July 2003. doi: 10.1029/2001jb001744.

De Pascale, G. P. and Langridge, R. M. New on-fault evidence for a great earthquake in A.D. 1717, central Alpine fault, New Zealand. *Geology*, 40(9):791–794, June 2012. doi: 10.1130/g33363.1.

DeMets, C. and Dixon, T. H. New kinematic models for Pacific-North America motion from 3 Ma to present, I: Evidence for steady motion and biases in the NUVEL-1A Model. *Geophysical Research Letters*, 26(13):1921–1924, July 1999. doi: 10.1029/1999gl900405.

DeVries, P. M. R., Krastev, P. G., Dolan, J. F., and Meade, B. J. Viscoelastic Block Models of the North Anatolian Fault: A Unified Earthquake Cycle Representation of Pre- and Postseismic Geodetic Observations. *Bulletin of the Seismological Society of America*, 107(1):403–417, Nov. 2016. doi: 10.1785/0120160059.

Dolan, J., Van Dissen, R., Rhodes, E., Zinke, R., Hatem, A., McGuire, C., Langridge, R., and Grenader, J. One tune, many tempos: Faults trade off slip in time and space to accommodate relative plate motions. *Earth and Planetary Science Letters*, 625:118484, Jan. 2024. doi: 10.1016/j.epsl.2023.118484.

Dolan, J. F. and Meade, B. J. A Comparison of Geodetic and Geologic Rates Prior to Large Strike-Slip Earthquakes: A Diversity of Earthquake-Cycle Behaviors? *Geochemistry, Geophysics, Geosystems*, 18(12):4426–4436, Dec. 2017. doi: 10.1002/2017gc007014.

Dolan, J. F., Bowman, D. D., and Sammis, C. G. Long-range and long-term fault interactions in Southern California. *Geology*, 35 (9):855, 2007. doi: 10.1130/g23789a.1.

Dolan, J. F., McAuliffe, L. J., Rhodes, E. J., McGill, S. F., and Zinke, R. Extreme multi-millennial slip rate variations on the Garlock fault, California: Strain super-cycles, potentially time-variable fault strength, and implications for system-level earthquake occurrence. *Earth and Planetary Science Letters*, 446:123–136, July 2016. doi: 10.1016/j.epsl.2016.04.011.

Elliott, J. and Freymueller, J. T. A Block Model of Present-Day Kinematics of Alaska and Western Canada. *Journal of Geophysical Research: Solid Earth*, 125(7):e2019JB018378, 2020. doi: 10.1029/2019JB018378.

Ellis, S. and Stöckhert, B. Elevated stresses and creep rates beneath the brittle-ductile transition caused by seismic faulting in the upper crust. *Journal of Geophysical Research: Solid Earth*, 109(B5), May 2004. doi: 10.1029/2003jb002744.

Evans, E. L. Persistent slip rate discrepancies in the eastern California (USA) shear zone: Reply. *Geology*, 45(9):e426–e426, Sept. 2017a. doi: 10.1130/g39439y.1.

Evans, E. L. A Comprehensive Analysis of Geodetic Slip-Rate Estimates and Uncertainties in California. *Bulletin of the Seismological Society of America*, 108(1):1–18, Nov. 2017b. doi: 10.1785/0120170159.

Evans, E. L., Thatcher, W. R., Pollitz, F. F., and Murray, J. R. Persistent slip rate discrepancies in the eastern California (USA) shear zone. *Geology*, 44(9):691–694, July 2016. doi: 10.1130/g37967.1.

Fougere, D., Dolan, J. F., Rhodes, E., McGill, S. F., and Ivester, A. Deciphering Non-Constant Earthquake Behavior: Insights from the Garlock Fault in Southern California. 2023. https://agu.confex.com/agu/fm23/meetingapp.cgi/Paper/1283318.

Fougere, D., Dolan, J. F., Rhodes, E. J., and McGill, S. F. Refined Holocene slip rate for the western and central segments of the Garlock fault: Record of Alternating Millenial-Scale Periods of Fast and Slow Fault Slip, submitted.

Friedrich, A. M., Wernicke, B. P., Niemi, N. A., Bennett, R. A., and Davis, J. L. Comparison of geodetic and geologic data from the Wasatch region, Utah, and implications for the spectral character of Earth deformation at periods of 10 to 10 million years. *Journal of Geophysical Research: Solid Earth*, 108(B4), Apr. 2003. doi: 10.1029/2001jb000682.

Ganev, P. N., Dolan, J. F., Blisniuk, K., Oskin, M., and Owen, L. A. Paleoseismologic evidence for multiple Holocene earthquakes on the Calico fault: Implications for earthquake clustering in the Eastern California shear zone. *Lithosphere*, 2(4):287–298, Aug. 2010. doi: 10.1130/l82.1.

Gauriau, J. and Dolan, J. F. Relative Structural Complexity of Plate-Boundary Fault Systems Controls Incremental Slip-Rate Behavior of Major Strike-Slip Faults. *Geochemistry, Geophysics, Geosystems*, 22(11), Nov. 2021. doi: 10.1029/2021gc009938.

Gauriau, J., Barbot, S., and Dolan, J. F. Islands of chaos in a sea of periodic earthquakes. *Earth and Planetary Science Letters*, 618: 118274, Sept. 2023. doi: 10.1016/j.epsl.2023.118274.

Gomez, F., Meghraoui, M., Darkal, A., Hijazi, F., Mouty, M., Suleiman, Y., Sbeinati, R., Darawcheh, R., Al-Ghazzi, R., and Barazangi, M. Holocene faulting and earthquake recurrence along the Serghaya branch of the Dead Sea fault system in Syria and Lebanon. *Geophys J Int*, 153(3):658–674, June 2003. doi: 10.1046/j.1365-246X.2003.01933.x.

Gomez, F., Nemer, T., Tabet, C., Khawlie, M., Meghraoui, M., and Barazangi, M. Strain partitioning of active transpression within the Lebanese restraining bend of the Dead Sea Fault (Lebanon and SW Syria). *Geological Society, London, Special Publications*, 290(1):285–303, Jan. 2007. doi: 10.1144/290.10.

Gomez, F., Cochran, W. J., Yassminh, R., Jaafar, R., Reilinger, R., Floyd, M., King, R. W., and Barazangi, M. Fragmentation of the Sinai Plate indicated by spatial variation in present-day slip rate along the Dead Sea Fault System. *Geophysical Journal International*, 221(3):1913–1940, Feb. 2020. doi: 10.1093/gji/ggaa095.

Grant Ludwig, L., Akciz, S. O., Arrowsmith, J. R., and Salisbury, J. B. Reproducibility of San Andreas Fault Slip Rate Measurements at Wallace Creek in the Carrizo Plain, CA. *Earth and Space Science*, 6(1):156–165, Jan. 2019. doi: 10.1029/2017ea000360.

Güvercin, S. E., Karabulut, H., Konca, A. O., Doğan, U., and Ergintav, S. Active seismotectonics of the East Anatolian Fault. *Geophysical Journal International*, 230(1):50–69, July 2022. doi: 10.1093/gji/ggac045.

Haddon, E. K., Amos, C. B., Zielke, O., Jayko, A. S., and Bürgmann, R. Surface slip during large Owens Valley earthquakes. *Geochemistry, Geophysics, Geosystems*, 17(6):2239–2269, June 2016. doi: 10.1002/2015gc006033.

Haeussler, P. J. Surface Rupture and Slip Distribution of the Denali and Totschunda Faults in the 3 November 2002 M 7.9 Earthquake, Alaska. *Bulletin of the Seismological Society of America*, 94(6B):S23–S52, Dec. 2004. doi: 10.1785/0120040626.

Haibing, L., Van der Woerd, J., Tapponnier, P., Klinger, Y., Xuexiang, Q., Jingsui, Y., and Yintang, Z. Slip rate on the Kunlun fault at Hongshui Gou, and recurrence time of great events comparable to the 14/11/2001, Mw7.9 Kokoxili earthquake. *Earth and Planetary Science Letters*, 237(1):285–299, Aug. 2005. doi: 10.1016/j.epsl.2005.05.041.

Hamiel, Y., Piatibratova, O., Mizrahi, Y., Nahmias, Y., and Sagy, A. Crustal Deformation across the Jericho Valley Section of the Dead Sea Fault as Resolved by Detailed Field and Geodetic Observations. *Geophysical Research Letters*, 45(7):3043–3050, 2018. doi: 10.1002/2018GL077547.

Hamling, I. J., Hreinsdóttir, S., Clark, K., Elliott, J., Liang, C., Fielding, E., Litchfield, N., Villamor, P., Wallace, L., Wright, T. J., D'Anastasio, E., Bannister, S., Burbidge, D., Denys, P., Gentle, P., Howarth, J., Mueller, C., Palmer, N., Pearson, C., Power, W., Barnes, P., Barrell, D. J. A., Van Dissen, R., Langridge, R., Little, T., Nicol, A., Pettinga, J., Rowland, J., and Stirling, M. Complex multifault rupture during the 2016 M w 7.8 Kaikōura earthquake, New Zealand. *Science*, 356(6334), Apr. 2017. doi: 10.1126/science.aam7194.

Handy, M. R. Deformation regimes and the rheological evolution of fault zones in the lithosphere: the effects of pressure, temperature, grainsize and time. *Tectonophysics*, 163(1–2):119–152, June 1989. doi: 10.1016/0040-1951(89)90122-4.

Handy, M. R., Hirth, G., and Bürgmann, R. Continental Fault Structure and Rheology from the Frictional-to-Viscous Transition Downward. In *Tectonic Faults*, page 139–182. The MIT Press, May 2007. doi: 10.7551/mitpress/6703.003.0008.

Hatem, A. E. and Dolan, J. F. A Model for the Initiation, Evolution, and Controls on Seismic Behavior of the Garlock Fault, California. *Geochemistry, Geophysics, Geosystems*, 19(7):2166–2178, July 2018. doi: 10.1029/2017gc007349.

Hatem, A. E., Dolan, J. F., Zinke, R. W., Van Dissen, R. J., McGuire, C. M., and Rhodes, E. J. A 2000 Yr Paleoearthquake Record along the Conway Segment of the Hope Fault: Implications for Patterns of Earthquake Occurrence in Northern South Island and Southern North Island, New Zealand. *Bulletin of the Seismological Society of America*, 109(6):2216–2239, Sept. 2019. doi: 10.1785/0120180313.

Hatem, A. E., Dolan, J. F., Zinke, R. W., Langridge, R. M., McGuire, C. P., Rhodes, E. J., Brown, N., and Van Dissen, R. J. Holocene to latest Pleistocene incremental slip rates from the east-central Hope fault (Conway segment) at Hossack Station, Marlborough fault system, South Island, New Zealand: Towards a dated path of earthquake slip along a plate boundary fault. *Geosphere*, 16 (6):1558–1584, Oct. 2020. doi: 10.1130/ges02263.1.

He, J., Vernant, P., Chéry, J., Wang, W., Lu, S., Ku, W., Xia, W., and Bilham, R. Nailing down the slip rate of the Altyn Tagh fault. *Geophysical Research Letters*, 40(20):5382–5386, 2013. doi: 10.1002/2013GL057497.

Hearn, E. "Ghost Transient" Corrections to the Southern California GPS Velocity Field from San Andreas Fault Seismic Cycle Models. *Seismological Research Letters*, 93(6):2973–2989, Aug. 2022. doi: 10.1785/0220220156.

Hearn, E. H., Pollitz, F. F., Thatcher, W. R., and Onishi, C. T. How do "ghost transients" from past earthquakes affect GPS slip rate estimates on southern California faults? *Geochemistry, Geophysics, Geosystems*, 14(4):828–838, Apr. 2013. doi: 10.1002/ggge.20080.

Hubert-Ferrari, A., Armijo, R., King, G., Meyer, B., and Barka, A. Morphology, displacement, and slip rates along the North Anatolian Fault, Turkey. *Journal of geophysical research Solid Earth*, 107 (B10):ETG 9–1–ETG 9–33, 2002. doi: 10.1029/2001JB000393.

Hussain, E., Wright, T. J., Walters, R. J., Bekaert, D. P. S., Lloyd, R., and Hooper, A. Constant strain accumulation rate between major earthquakes on the North Anatolian Fault. *Nature Communications*, 9(1), Apr. 2018. doi: 10.1038/s41467-018-03739-2.

Johnson, K. M., Wallace, L. M., Maurer, J., Hamling, I. J., Williams, C. A., Rollins, C., Gerstenberger, M. C., and Van Dissen, R. J. Geodetic deformation model for the 2022 update of the New Zealand National Seismic Hazard Model. 2022. doi: 10.21420/P93X-8293.

Kirby, E., Harkins, N., Wang, E., Shi, X., Fan, C., and Burbank, D. Slip rate gradients along the eastern Kunlun fault. *Tectonics*, 26(2), 2007. doi: 10.1029/2006TC002033.

Kirby, E., Anandakrishnan, S., Phillips, F., and Marrero, S. Late Pleistocene slip rate along the Owens Valley fault, eastern California. *Geophysical Research Letters*, 35(1), 2008. doi: 10.1029/2007GL031970.

Klinger, Y., Avouac, J., Abou Karaki, N., Dorbath, L., Bourles, D., and Reyss, J. Slip rate on the Dead Sea transform fault in northern Araba valley (Jordan). *Geophysical Journal International*, 142(3): 755–768, 2000. doi: 10.1046/j.1365-246x.2000.00165.x.

Kondo, H., Özaksoy, V., and Yıldirim, C. Slip history of the 1944 Bolu-Gerede earthquake rupture along the North Anatolian fault system: Implications for recurrence behavior of multisegment earthquakes. *Journal of Geophysical Research: Solid Earth*, 115(B4), Apr. 2010. doi: 10.1029/2009jb006413.

Kozacı, O., Dolan, J., Finkel, R., and Hartleb, R. Late Holocene slip rate for the North Anatolian fault, Turkey, from cosmogenic 36Cl geochronology: Implications for the constancy of fault loading and strain release rates. *Geological Society of America Bulletin*, 35(10):867–870, 2007.

Kozacı, O., Dolan, J. F., and Finkel, R. C. A late Holocene slip rate for the central North Anatolian fault, at Tahtaköprü, Turkey, from cosmogenic 10Be geochronology: Implications for fault loading and strain release rates. *Journal of Geophysical Research: Solid Earth*, 114(B1), Jan. 2009. doi: 10.1029/2008jb005760.

Kozacı, O., Dolan, J. F., Yönlü, O., and Hartleb, R. D. Paleoseismologic evidence for the relatively regular recurrence of infrequent, large-magnitude earthquakes on the eastern North Anatolian fault at Yaylabeli, Turkey. *Lithosphere*, 3(1):37–54, Feb. 2011. doi: 10.1130/l118.1.

Kurt, H., Sorlien, C. C., Seeber, L., Steckler, M. S., Shillington, D. J., Cifci, G., Cormier, M., Dessa, J., Atgin, O., Dondurur, D., Demirbag, E., Okay, S., Imren, C., Gurcay, S., and Carton, H. Steady late quaternary slip rate on the Cinarcik section of the North Anatolian fault near Istanbul, Turkey. *Geophysical Research Letters*, 40(17):4555–4559, Sept. 2013. doi: 10.1002/grl.50882.

Li, C., Zhang, P.-z., Yin, J., and Min, W. Late Quaternary left-lateral slip rate of the Haiyuan fault, northeastern margin of the Tibetan Plateau. *Tectonics*, 28(5), 2009. doi: 10.1029/2008TC002302.

Liu, J., Ren, Z., Zhang, H., Li, C., Zhang, Z., WenJun, Z., XueMei, L. I., and CaiCai, L. I. U. Late Quaternary slip rate of the Laohushan fault within the Haiyuan fault zone and its tectonic implications. *Chinese Journal of Geophysics*, 61(4):1281–1297, Apr. 2018. doi: 10.6038/cjg2018L0364.

Liu, J., Ren, Z., Zhang, H., Li, C., Zhang, Z., Zheng, W., Li, X., and Liu, C. Slip Rates Along the Laohushan Fault and Spatial Variation in Slip Rate Along the Haiyuan Fault Zone. *Tectonics*, 41(2): e2021TC006992, 2022. doi: 10.1029/2021TC006992.

Luo, G. and Liu, M. Stress evolution and fault interactions before and after the 2008 Great Wenchuan earthquake. *Tectonophysics*, 491(1–4):127–140, Aug. 2010. doi: 10.1016/j.tecto.2009.12.019.

Mancktelow, N. S. and Pennacchioni, G. The influence of grain boundary fluids on the microstructure of quartz-feldspar mylonites. *Journal of Structural Geology*, 26(1):47–69, Jan. 2004. doi: 10.1016/s0191-8141(03)00081-6.

Masson, F., Hamiel, Y., Agnon, A., Klinger, Y., and Deprez, A. Variable behavior of the Dead Sea Fault along the southern Arava segment from GPS measurements. *Comptes Rendus Geoscience*, 347(4):161–169, July 2015. doi: 10.1016/j.crte.2014.11.001.

Matmon, A., Schwartz, D., Haeussler, P., Finkel, R., Lienkaemper, J., Stenner, H., and Dawson, T. Denali fault slip rates and Holocene–late Pleistocene kinematics of central Alaska. *Geology*, 34(8):645–648, Aug. 2006. doi: 10.1130/G22361.1.

McGill, S. and Sieh, K. Holocene slip rate of the Central Garlock Fault in southeastern Searles Valley, California. *Journal of Geophysical Research: Solid Earth*, 98(B8):14217–14231, 1993. doi: 10.1029/93JB00442.

Meade, B. J., Klinger, Y., and Hetland, E. A. Inference of Multiple Earthquake-Cycle Relaxation Timescales from Irregular Geodetic Sampling of Interseismic Deformation. *Bulletin of the Seismological Society of America*, 103(5):2824–2835, Sept. 2013. doi: 10.1785/0120130006.

Meghraoui, M., Aksoy, M. E., Akyüz, H. S., Ferry, M., Dikbaş, A., and Altunel, E. Paleoseismology of the North Anatolian Fault at Güzelköy (Ganos segment, Turkey): Size and recurrence time of earthquake ruptures west of the Sea of Marmara. *Geochemistry, Geophysics, Geosystems*, 13(4), Apr. 2012. doi: 10.1029/2011gc003960.

Melosh, B. L., Rowe, C. D., Gerbi, C., Smit, L., and Macey, P. Seismic cycle feedbacks in a mid-crustal shear zone. *Journal of Structural Geology*, 112:95–111, July 2018. doi: 10.1016/j.jsg.2018.04.004.

Mencin, D., Bendick, R., Upreti, B. N., Adhikari, D. P., Gajurel, A., Bhattarai, R. R., Shrestha, H. R., Bhattarai, T. N., Manandhar, N., Galetzka, J., Knappe, E., Pratt-Sitaula, B., Aoudia, A., and Bilham, R. Himalayan strain reservoir inferred from limited afterslip following the Gorkha earthquake. *Nature Geoscience*, 9(7): 533–537, June 2016. doi: 10.1038/ngeo2734.

Mildon, Z. K., Roberts, G. P., Faure Walker, J. P., Beck, J., Papanikolaou, I., Michetti, A. M., Toda, S., Iezzi, F., Campbell, L., McCaffrey, K. J. W., Shanks, R., Sgambato, C., Robertson, J., Meschis, M., and Vittori, E. Surface faulting earthquake clustering controlled by fault and shear-zone interactions. *Nature Communications*, 13(1), Nov. 2022. doi: 10.1038/s41467-022-34821-5.

Nicol, A. and Dissen, R. V. A 6000-year record of surface-rupturing paleoearthquakes on the Wairau Fault, New Zealand. *New Zealand Journal of Geology and Geophysics*, 61(3):341–358, July 2018. doi: 10.1080/00288306.2018.1498360.

Nicol, A., Howell, A., Litchfield, N., Wilson, T., Bannister, S., and Massey, C. Introduction to the Kaikōura earthquake special issue. *New Zealand Journal of Geology and Geophysics*, 66(2): 137–146, Apr. 2023. doi: 10.1080/00288306.2023.2197240.

Niemi, T. M., Zhang, H., Atallah, M., and Harrison, J. B. J. Late Pleistocene and Holocene slip rate of the Northern Wadi Araba fault, Dead Sea Transform, Jordan. *Journal of Seismology*, 5(3): 449–474, July 2001. doi: 10.1023/A:1011487912054.

Noriega, G., Arrowsmith, J., Grant, L., and Young, J. Stream Channel Offset and Late Holocene Slip Rate of the San Andreas Fault at the Van Matre Ranch Site, Carrizo Plain, California. *Bulletin of the Seismological Society of America*, 96(1):33–47, 2006. doi: 10.1785/0120050094.

Okazaki, K., Burdette, E., and Hirth, G. Rheology of the Fluid Oversaturated Fault Zones at the Brittle-Plastic Transition. *Journal of Geophysical Research: Solid Earth*, 126(2), Feb. 2021. doi: 10.1029/2020jb020804.

Okudaira, T., Shigematsu, N., Harigane, Y., and Yoshida, K. Grain size reduction due to fracturing and subsequent grain-size-sensitive creep in a lower crustal shear zone in the presence of a CO2-bearing fluid. *Journal of Structural Geology*, 95:171–187, Feb. 2017. doi: 10.1016/j.jsg.2016.11.001.

Onderdonk, N. W., McGill, S. F., and Rockwell, T. K. Short-term variations in slip rate and size of prehistoric earthquakes during the past 2000 years on the northern San Jacinto fault zone, a major plate-boundary structure in southern California. *Lithosphere*, 7 (3):211–234, Feb. 2015. doi: 10.1130/l393.1.

Oskin, M., Perg, L., Blumentritt, D., Mukhopadhyay, S., and Iriondo, A. Slip rate of the Calico fault: Implications for anomalous geodetic strain accumulation across the Eastern California shear zone. 2004:G11A–0776, Dec. 2004. https://ui.adsabs.harvard.edu/abs/2004AGUFM.G11A0776O.

Oskin, M., Perg, L., Blumentritt, D., Mukhopadhyay, S., and Iriondo, A. Slip rate of the Calico fault: Implications for geologic versus geodetic rate discrepancy in the Eastern California Shear Zone. *Journal of Geophysical Research: Solid Earth*, 112(B3), Mar. 2007. doi: 10.1029/2006jb004451.

Oskin, M., Perg, L., Shelef, E., Strane, M., Gurney, E., Singer, B., and Zhang, X. Elevated shear zone loading rate during an earthquake cluster in eastern California. *Geology*, 36(6):507, 2008. doi: 10.1130/g24814a.1.

Page, C. J., Denys, P. H., and Pearson, C. F. A geodetic study of the Alpine Fault through South Westland: using campaign GPS data to model slip rates on the Alpine Fault. *New Zealand Journal of Geology and Geophysics*, 61(3):359–366, July 2018. doi: 10.1080/00288306.2018.1494006.

Peltzer, G., Crampé, F., Hensley, S., and Rosen, P. Transient strain accumulation and fault interaction in the Eastern California shear zone. *Geology*, 29(11):975, 2001. doi: 10.1130/0091-7613(2001)029<0975:tsaafi>2.0.co;2.

Reid, H. *The California Earthquake of April 18, 1906, Report of the State Earthquake Investigation Commission*. Carnegie Institution of Washington, Washington, D.C, 1910.

Reilinger, R., McClusky, S., Vernant, P., Lawrence, S., Ergintav, S., Cakmak, R., and Ozener, H. GPS constraints on continental deformation in the Africa-Arabia- Eurasia continental collision zone and implications for the dynamics of plate interactions. *Journal of geophysical research*, 111(B5), 2006. doi: 10.10292005JB004051.

Rockwell, T. K., Lindvall, S., Herzberg, M., Murbach, D., Dawson, T., and Berger, G. Paleoseismology of the Johnson Valley, Kickapoo, and Homestead Valley Faults: Clustering of Earthquakes in the Eastern California Shear Zone. *Bulletin of the Seismological Society of America*, 90(5):1200–1236, Oct. 2000. doi: 10.1785/0119990023.

Salisbury, J. B., Arrowsmith, J. R., Brown, N., Rockwell, T., Akciz, S., and Ludwig, L. G. The Age and Origin of Small Offsets at Van

Matre Ranch along the San Andreas Fault in the Carrizo Plain, California. *Bulletin of the Seismological Society of America*, 108 (2):639–653, Jan. 2018. doi: 10.1785/0120170162.

Scharer, K., Weldon, R., Biasi, G., Streig, A., and Fumal, T. Ground-rupturing earthquakes on the northern Big Bend of the San Andreas Fault, California, 800 A.D. to Present. *Journal of Geophysical Research: Solid Earth*, 122(3):2193–2218, Mar. 2017. doi: 10.1002/2016jb013606.

Shao, Y., Liu-Zeng, J., Van der Woerd, J., Klinger, Y., Oskin, M. E., Zhang, J., Wang, P., Wang, P., Wang, W., and Yao, W. Late Pleistocene slip rate of the central Haiyuan fault constrained from optically stimulated luminescence, 14C, and cosmogenic isotope dating and high-resolution topography. *GSA Bulletin*, 133 (7-8):1347–1369, Oct. 2020. doi: 10.1130/B35571.1.

Shen, Z.-K., Wang, M., Li, Y., Jackson, D. D., Yin, A., Dong, D., and Fang, P. Crustal deformation along the Altyn Tagh fault system, western China, from GPS. *Journal of Geophysical Research: Solid Earth*, 106(B12):30607–30621, 2001. doi: 10.1029/2001JB000349.

Shimazaki, K. and Nakata, T. Time-predictable recurrence model for large earthquakes. *Geophysical Research Letters*, 7(4): 279–282, Apr. 1980. doi: 10.1029/gl007i004p00279.

Sieh, K. E. and Jahns, R. Holocene activity of the San Andreas fault at Wallace Creek, California. *Geological Society of America Bulletin*, 95:883–896, 1984. doi: 10.1130/0016-7606(1984)95<883:HAOTSA>2.0.CO;2.

Sutherland, R., Berryman, K., and Norris, R. Quaternary slip rate and geomorphology of the Alpine fault: Implications for kinematics and seismic hazard in southwest New Zealand. *Geological Society of America Bulletin*, 118(3–4):464–474, Mar. 2006. doi: 10.1130/b25627.1.

Tong, X., Smith-Konter, B., and Sandwell, D. T. Is there a discrepancy between geological and geodetic slip rates along the San Andreas Fault System? *Journal of Geophysical Research: Solid Earth*, 119(3):2518–2538, Mar. 2014. doi: 10.1002/2013jb010765.

Ulrich, T., Gabriel, A.-A., Ampuero, J.-P., and Xu, W. Dynamic viability of the 2016 Mw 7.8 Kaikōura earthquake cascade on weak crustal faults. *Nature Communications*, 10(1), Mar. 2019. doi: 10.1038/s41467-019-09125-w.

Van Der Woerd, J., Tapponnier, P., J. Ryerson, F., Meriaux, A.-S., Meyer, B., Gaudemer, Y., Finkel, R. C., Caffee, M. W., Guoguan, Z., and Zhiqin, X. Uniform postglacial slip-rate along the central 600 km of the Kunlun Fault (Tibet), from 26Al, 10Be, and 14C dating of riser offsets, and climatic origin of the regional morphology. *Geophysical Journal International*, 148(3):356–388, Mar. 2002. doi: 10.1046/j.1365-246x.2002.01556.x.

van Dissen, R. Slip rate variations on major strike-slip faults in central New Zealand and potential impacts on hazard estimation, 2020. https://repo.nzsee.org.nz/handle/nzsee/1691.

Wallace, L. M., Barnes, P., Beavan, J., Dissen, R. V., Litchfield, N., Mountjoy, J., Langridge, R., Lamarche, G., and Pondard, N. The kinematics of a transition from subduction to strike-slip: An example from the central New Zealand plate boundary. *Journal of Geophysical Research: Solid Earth*, 117(B2), 2012. doi: https://doi.org/10.1029/2011JB008640.

Wechsler, N., Rockwell, T. K., and Klinger, Y. Variable slip-rate and slip-per-event on a plate boundary fault: The Dead Sea fault in northern Israel. *Tectonophysics*, 722:210–226, Jan. 2018. doi: 10.1016/j.tecto.2017.10.017.

Weldon, R., Scharer, K., Fumal, T., and Biasi, G. Wrightwood and the earthquake cycle: What a long recurrence record tells us about how faults work. *GSA Today*, 14(9):4, 2004. doi: 10.1130/1052-5173(2004)014<4:watecw>2.0.co;2.

Weldon, R. J. and Sieh, K. E. Holocene rate of slip and tentative recurrence interval for large earthquakes on the San Andreas fault, Cajon Pass, southern California. *Geological Society of America Bulletin*, 96(6):793, 1985. doi: 10.1130/0016-7606(1985)96<793:hrosat>2.0.co;2.

Yönlü, O. and Karabacak, V. Surface rupture history and 18 kyr long slip rate along the Pazarcık segment of the East Anatolian Fault. *Journal of the Geological Society*, 181(1), Dec. 2023. doi: 10.1144/jgs2023-056.

Zhang, P.-Z., Molnar, P., and Xu, X. Late Quaternary and present-day rates of slip along the Altyn Tagh Fault, northern margin of the Tibetan Plateau. *Tectonics*, 26(5), 2007. doi: 10.1029/2006TC002014.

Zhao, D., Qu, C., Bürgmann, R., Gong, W., Shan, X., Qiao, X., Zhao, L., Chen, H., and Liu, L. Large-Scale Crustal Deformation, Slip-Rate Variation, and Strain Distribution Along the Kunlun Fault (Tibet) From Sentinel-1 InSAR Observations (2015–2020). *Journal of Geophysical Research: Solid Earth*, 127 (1):e2021JB022892, 2022. doi: 10.1029/2021JB022892.

Zinke, R., Dolan, J. F., Rhodes, E. J., Van Dissen, R., and McGuire, C. P. Highly Variable Latest Pleistocene-Holocene Incremental Slip Rates on the Awatere Fault at Saxton River, South Island, New Zealand, Revealed by Lidar Mapping and Luminescence Dating. *Geophysical Research Letters*, 44(22), Nov. 2017. doi: 10.1002/2017gl075048.

Zinke, R., Dolan, J. F., Rhodes, E. J., Van Dissen, R., McGuire, C. P., Hatem, A. E., Brown, N. D., and Langridge, R. M. Multimillennial Incremental Slip Rate Variability of the Clarence Fault at the Tophouse Road Site, Marlborough Fault System, New Zealand. *Geophysical Research Letters*, 46(2):717–725, Jan. 2019. doi: 10.1029/2018gl080688.

Zinke, R., Dolan, J. F., Rhodes, E. J., Van Dissen, R. J., Hatem, A. E., McGuire, C. P., Brown, N. D., and Grenader, J. R. Latest Pleistocene–Holocene Incremental Slip Rates of the Wairau Fault: Implications for Long-Distance and Long-Term Coordination of Faulting Between North and South Island, New Zealand. *Geochemistry, Geophysics, Geosystems*, 22(9), Sept. 2021. doi: 10.1029/2021gc009656.

doi:10.26443/seismica.v3i1.1098

SEISMICA

Shear-wave attenuation anisotropy: a new constraint on mantle melt near the Main Ethiopian Rift

Joseph Asplet * [1,2], **James Wookey** [1], **J-Michael Kendall** [2], **Mark Chapman** [3], **Ritima Das** [4]

[1] School of Earth Sciences, University of Bristol, Bristol, U.K., [2] Department of Earth Science, University of Oxford, Oxford, U.K., [3] School of Geosciences, University of Edinburgh, Edinburgh, Scotland, [4] Department of Earth Sciences, Pondicherry University, Pondicherry, India

Author contributions: *Conceptualization*: J Asplet, J Wookey, M Kendall. *Methodology*: J Asplet, J Wookey, R Das. *Software*: J Asplet, J Wookey. *Formal Analysis*: J Asplet, J Wookey, J-M Kendall. *Writing - original draft*: J Asplet. *Writing - Review & Editing*: J Asplet, J Wookey, J-M Kendall, M Chapman. *Visualization*: J Asplet. *Funding acquisition*: J Wookey, J-M Kendall.

Production Editor: Gareth Funning
Handling Editor: Suzan van der Lee
Copy & Layout Editor: Hannah F. Mark

Signed reviewer(s): Robert Nowack, Rasheed Ajala

Received: August 23, 2023
Accepted: April 21, 2024
Published: May 13, 2024

Abstract The behaviour of fluids in preferentially aligned fractures plays an important role in a range of dynamic processes within the Earth. In the near-surface, understanding systems of fluid-filled fractures is crucial for applications such as geothermal energy production, monitoring CO2 storage sites, and exploration for metalliferous sub-volcanic brines. Mantle melting is a key geodynamic process, exerting control over its composition and dynamic processes. Upper mantle melting weakens the lithosphere, facilitating rifting and other surface expressions of tectonic processes. Aligned fluid-filled fractures are an efficient mechanism for seismic velocity anisotropy, requiring very low volume fractions, but such rock physics models also predict significant shear-wave attenuation anisotropy. In comparison, the attenuation anisotropy expected for crystal preferred orietation mechanisms is negligible or would only operate outside of the seismic frequency band. Here we demonstrate a new method for measuring shear-wave attenuation anisotropy, apply it to synthetic examples, and make the first measurements of SKS attenuation anisotropy using data recorded at the station FURI, in Ethiopia. At FURI we measure attenuation anisotropy where the fast shear-wave has been more attenuated than the slow shear-wave. This can be explained by the presence of aligned fluids, most probably melts, in the upper mantle using a poroelastic squirt flow model. Modelling of this result suggests that a 1% melt fraction, hosted in aligned fractures dipping ca. 40° that strike perpendicular to the Main Ethiopian Rift, is required to explain the observed attenuation anisotropy. This agrees with previous SKS shear-wave splitting analysis which suggested a 1% melt fraction beneath FURI. The interpreted fracture strike and dip, however, disagrees with previous work in the region which interprets sub-vertical melt inclusions aligned parallel to the Main Ethiopian Rift which only produce attenuation anisotropy where the slow shear-wave is more attenuated. These results show that attenuation anisotropy could be a useful tool for detecting mantle melt, and may offer strong constraints on the extent and orientation of melt inclusions which cannot be achieved from seismic velocity anisotropy alone.

Non-technical summary When seismic signals travel through the Earth they lose energy, or attenuate, due to various mechanisms including the nature of the rocks they propagate through. One particularly strong mechanism is the presence of fluids, such as water or molten rock, in pore spaces. Theory from rock mechanics predicts that if fluids are hosted in aligned fractures then the loss of energy depends on the propagation direction of the earthquake signal. This predicts a difference in the loss of energy between two coupled shear-waves. Measuring this difference in energy loss then would give us a powerful tool to detect and quantify the presence of fluids in the subsurface. Here we describe a new method to measure this difference in energy loss between two shear-waves by measuring a difference in frequency content. We demonstrate this method for synthetic seismic signals, and make the first measurements for teleseismic shear-wave data. We use seismic waves that sample the upper mantle beneath the seismic station FURI, which is situated near Addis Ababa, Ethiopia. We find that our new observations can be explained by a 1% volume fraction of molten material, which agrees with previous interpretations made for FURI. Modelling using current rock physics models suggests that this requires aligned fractures that dip 40° and are oriented perpendicular to the Main Ethiopian Rift which disagrees with previous interpretations of melt orientated parallel to the Main Ethiopian Rift.

1 Introduction

The presence of fluids within a fractured host rock has important effects on its seismic and mechanical properties. In the crust, there are many systems where the presence of fluids is critical. These include melt-water pockets in glaciers, hydrocarbons in fractured reservoirs, and hydrothermal and magmatic systems beneath volcanos. Melting is also a key process within the mantle, exerting control over mantle composition and dynamic processes. Upper mantle melt weakens the lithosphere, facilitating rifting (e.g., Buck, 2004; Kendall et al., 2005) and other surface expressions of tectonic

*Corresponding author: joseph.asplet@bristol.ac.uk

processes. Observed low seismic velocity zones in the mantle transition zone (e.g., Schmandt et al., 2014; Liu et al., 2016b) and ultra-low velocity zones (ULVZs, e.g., Liu et al., 2016a; Li et al., 2022) in the lowermost mantle have been interpreted in terms of melt. Aligned melt pockets are a very efficient mechanism for generating seismic anisotropy (e.g., Kendall, 2000; Holtzman and Kendall, 2010). This makes it difficult to discriminate between melt, other shape-preferred orientation (such as dry cracks), and lattice-preferred orientation models of seismic anisotropy from the crust (e.g., Bacon et al., 2022) to the lowermost mantle (e.g., Asplet et al., 2022).

Rock physics models predict that aligned sets of fluid-filled fractures, or melt inclusions, produce an effective medium that exhibits both velocity and attenuation anisotropy (e.g., Hudson, 1980; Chapman, 2003; Jin et al., 2018). This result can be achieved by either treating cracks as scatterers (Figure 1a,b; Hudson, 1980) or through the poroelastic squirt flow of fluids in saturated (or partially saturated) meso-scale fractures (Figure 1d,e; Chapman, 2003; Galvin and Gurevich, 2009; Rubino and Holliger, 2012; Jin et al., 2018; Solazzi et al., 2021). The squirt flow model, in particular, predicts a strong dependence of attenuation anisotropy on the presence of fluids (such as melt) and fracture properties.

Whilst attenuation anisotropy can be observed for P waves (e.g., Liu et al., 2007; Ford et al., 2022) it is the attenuation of S-waves that interests us here. Both the crack scattering (Figure 1b) and squirt flow (Figure 1e) models predict an attenuation anisotropy which can be used to complement studies that measure velocity anisotropy using shear-wave splitting (e.g. Kendall et al., 2005; Verdon and Kendall, 2011; Al-Harrasi et al., 2011; Baird et al., 2013, 2015; Bacon et al., 2022; Schlaphorst et al., 2022). Attenuation anisotropy is a highly sensitive tool for detecting fluids within the earth that are hosted within aligned fractures. For microseismic settings, where the mechanism of seismic anisotropy is known to be fluid-filled fractures, measurements of anisotropic attenuation in shear-waves have been used to help constrain fracture and fluid properties (Carter and Kendall, 2006; Usher et al., 2017). Attenuation anisotropy can be observed directly in experiments (e.g., Best et al., 2007; Zhubayev et al., 2016), albeit at higher frequencies. Numerical models also show that attenuation anisotropy is sensitive to fluid transport properties (Wenzlau et al., 2010).

Measurements of differential attenuation between different teleseismic shear-wave phases, typically S-ScS, have been previously used to measure isotropic Q_S in the Earth's mantle (e.g., Lawrence and Wysession, 2006; Ford et al., 2012; Durand et al., 2013; Liu and Grand, 2018). This differential attenuation can be measured by either taking log-spectral ratios or by measuring instantaneous frequency relative to a reference seismogram (Matheney and Nowack, 1995). Here we employ an instantaneous frequency method, which has been shown to be more robust than spectral ratios for teleseismic shear-waves (Ford et al., 2012; Durand et al., 2013). By making measurements of differential attenuation between fast and slow split shear-waves it is possible to measure attenuation anisotropy. As attenuation anisotropy is primarily predicted by effective medium models of fluid-filled fractures, these measurements are highly sensitive to the presence of fluids, such as melt, within the Earth.

We outline how an instantaneous frequency matching method can be applied to measure attenuation anisotropy using shear-wave splitting. Using synthetic shear-wave data we demonstrate the frequency domain effects of attenuation anisotropy and the implications this can have for measurements of shear-wave splitting. We explore the pitfalls of measuring attenuation anisotropy and demonstrate the efficacy of our instantaneous frequency-matching method. We then demonstrate the application of joint measurements of attenuation anisotropy and shear-wave splitting using SKS data recorded at FURI, Ethiopia.

2 Models of attenuation anisotropy

When a shear-wave propagates through an anisotropic medium, seismic birefringence — or shear-wave splitting — occurs. The fast and slow shear-waves are polarised along the fast velocity direction and an (assumed) orthogonal direction and propagate at different velocities through the medium. This introduces a time delay between the two and can decouple the two (quasi) shear-waves, although in the teleseismic case the time delay time, δt, is much less than the dominant period of the waveform. Assuming that the medium can be described by a single elastic tensor c_{ijkl} the phase velocities and polarisation of each wave can be found by solving the Christoffel equation,

$$(c_{ijkl}n_j n_l - \rho V^2 \delta_{ik})p_k = 0\,, \tag{1}$$

where V is phase velocity, ρ is density, p_k is polarisation unit vector and $n_{j,l}$ are propagation unit vectors. Solving this eigenproblem yields three positive, real eigenvalues corresponding to $\rho V_P, \rho V_{S1}, \rho V_{S2}$ with corresponding eigenvectors describing the polarisation directions, which are mutually perpendicular (Mainprice, 2015).

If the medium is also attenuating, then both shear-waves experience a frequency-dependent loss in amplitude and dispersion. The isotropic attenuation of a shear-wave over its path length, l, can be described by the anelastic delay time t^* which is given by

$$t^* = \int_{\text{path}} \frac{dl}{v_S Q_S}\,, \tag{2}$$

where v_S is the isotropic shear-wave velocity and $1/Q_S$ is the isotropic shear-wave dissipation coefficient. It can be shown that an attenuating medium requires frequency-dependent velocities, or physical dispersion, where the intrinsic seismic velocity of waves propagating through a medium varies with frequency (Aki and Richards, 1980). If this physical dispersion is also anisotropic, then the seismic velocity anisotropy is frequency-dependent and it follows that attenuation is anisotropic also (Carter and Kendall, 2006).

In the case of an anisotropic attenuating medium, where the shear-wave dissipation coefficient $1/Q_S$ varies with propagation direction, the fast and slow split shear-waves will experience different anelastic delay times. We define this difference in anelastic delay times as

$$\Delta t^* = t^*_{S2} - t^*_{S1}, = \frac{l}{v_{S2}Q_{S2}} - \frac{l}{v_{S1}Q_{S1}}, \quad (3)$$

where $S1$ is the fast split shear-wave and $S2$ is the slow split shear-wave. Following this definition, a positive Δt^* represents the case where the slow shear-wave is more attenuated than the fast shear-wave and a negative Δt^* is where the fast shear-wave is more attenuated that the slow shear-wave. It is also worth noting that due to the definition of anelastic delay time (3) velocity anisotropy will produce a Δt^* even if there is isotropic attenuation (i.e., where $Q_{S2} = Q_{S1}$). This effect, however, due to the difference in travel times through the attenuating medium, is negligible compared to the Δt^* that can be predicted for anisotropic attenuation and will always produce $\Delta t^* > 0$ (Supplemental Figure 1).

2.1 Anisotropic attenuation due to fluid-filled fractures

We consider two main models of seismic anisotropy due to fluid-filled fractures which also allow for the modelling of attenuation anisotropy. These model attenuation due to scattering (Hudson, 1980) and due to poroelastic squirt flow of the hosted fluids (Chapman, 2003). Hudson (1980) employs an effective medium approach to model attenuation due to preferential scattering by the aligned fractures. For this reason, we refer to this model as crack scattering or simply scattering. The attenuation predicted by this model is anisotropic and frequency-dependent (e.g., Crampin, 1984). Crack scattering also predicts anisotropic attenuation for unsaturated (or dry) aligned cracks, although the attenuation profiles are sufficiently different to allow the dry and saturated cases to be distinguished (Crampin, 1984). The thin layering of material could also produce an effective medium with frequency-dependent anisotropy (Backus, 1962; Werner and Shapiro, 1999) and therefore attenuation anisotropy through a similar scattering mechanism.

There are, however, several limitations to this effective medium approach. It does not model the frequency dependence of the elastic constants, limiting the sensitivity to fracture size, and it neglects the effects of fluid exchange between fractures or between fractures and the host rock matrix. Work to extend the models to include such fluid interchange and equant porosity in the rock matrix show that this has a significant effect on the predicted seismic anisotropy (e.g., Thomsen, 1995; Hudson et al., 1996; Tod, 2001). To adequately model this system an approach that considers the poroelastic squirt flow of fluids held in a random collection of grain-scale microcracks and spherical pores along with aligned meso-scale fracture sets (i.e., fractures much larger than the grain scale) was developed (Chapman, 2003). In the poroelastic squirt flow model, the propagation of a seismic wave causes fluids to migrate between connected meso-scale fracture, micro-scale crack and pore spaces which results in frequency-dependent velocity and attenuation anisotropy. These poroelastic effects can also be modelled by treating the effect of pores and fractures as perturbations in an isotropic background medium (e.g., Jakobsen et al., 2003; Galvin and Gurevich, 2009, 2015). More recent developments in squirt flow models allow for partially saturated media (e.g., Rubino and Holliger, 2012; Solazzi et al., 2021) and for multi-phase fluids such as water and supercritical CO_2 (Jin et al., 2018). In both cases, squirt flow predicts attenuation anisotropy but we shall only consider the fully saturated case here. It should be noted that this model, and the scattering model, assume perfectly aligned fractures which is unlikely to represent real-world fracture systems completely. The models are also limited to very low aspect ratios which ultimately derives from the low aspect ratio limit of Eshelby's theory (Eshelby, 1957), which results in very low volume fractions (ca. 2×10^{-5}) of fluids required to be in aligned fracture to produce significant velocity and attenuation anisotropy. Recent numerical modelling of squirt flow dispersion models has shown that dispersion increases with fracture density and decreases with aspect ratio, with aspect ratios ≥ 0.1 showing very weak attenuation (Sun et al., 2020).

To calculate seismic velocity and attenuation anisotropy for both the crack scattering and squirt flow models we follow the approach of Crampin (1981). We can include attenuation in the definition of a medium's elastic tensor c_{ijkl} by introducing imaginary parts c^I_{ijkl} of complex elastic constants,

$$c_{ijkl} = c^R_{ijkl} + ic^I_{ijkl}\,, \quad (4)$$

where the real components c^R_{ijkl} are the elastic constants. Solving the Christoffel equation for this complex elastic tensor now yields complex eigenvalues $\lambda = \lambda^R + i\lambda^I$, with the dissipation coefficient $1/Q$ given by the ratio of the imaginary and real parts (Crampin, 1984),

$$\frac{1}{Q_P} = \frac{\lambda^I_P}{\lambda^R_P}\,, \quad (5)$$

$$\frac{1}{Q_{S1}} = \frac{\lambda^I_{S1}}{\lambda^R_{S1}}\,, \quad (6)$$

$$\frac{1}{Q_{S2}} = \frac{\lambda^I_{S2}}{\lambda^R_{S2}}\,. \quad (7)$$

For the crack scattering model, the imaginary components of the complex elastic tensor can be constructed using equations from Crampin (1984). Figure 1 shows the seismic velocity (Figure 1a) and attenuation (Figure 1b) profiles modelled as a function of propagation angle relative to the crack normal for a saturated, cracked solid for a frequency of $0.1\,\mathrm{Hz}$. The isotropic solid has velocities $v_p = 6.5\,\mathrm{km\,s^{-1}}$, $v_s = 3.6\,\mathrm{km\,s^{-1}}$ and a density of $2700\,\mathrm{kg\,m^{-3}}$ with fractures which are filled with a fluid with a P-wave velocity $v_p = 2.7\,\mathrm{km\,s^{-1}}$, a crack radius of $5\,\mathrm{km}$, a crack density of 0.05 and an aspect ratio of 1×10^{-4}. The predicted attenuation anisotropy, Δt^*, as a function of propagation angle (Figure 1c) is calculated using (3) assuming a path length of $50\,\mathrm{km}$ through

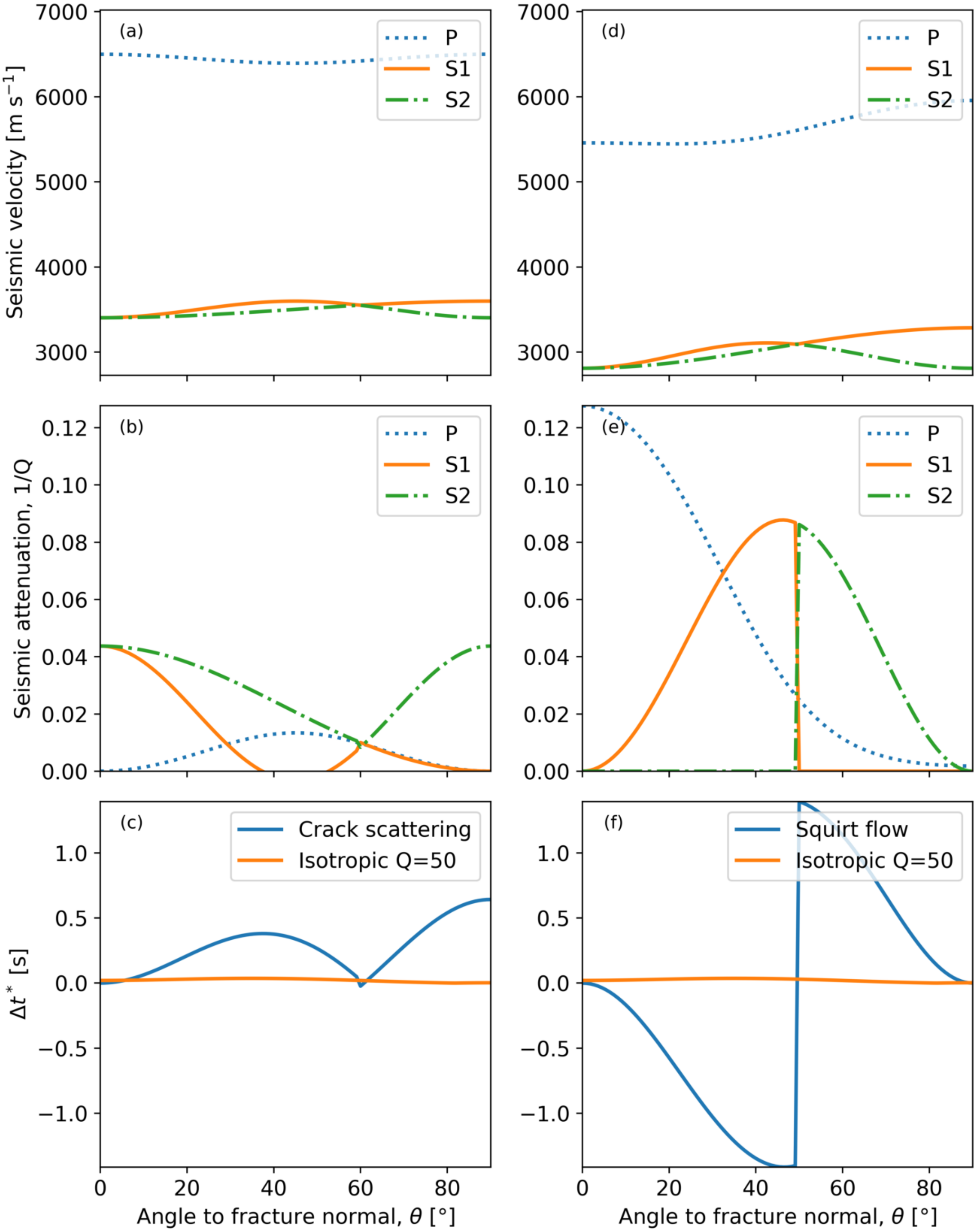


Figure 1 Seismic velocity, quality factor and attenuation anisotropy (expressed in terms of Δt^*) predicted by rock physics models for cracked, fluid-filled media which considers crack scattering (a,b,c Hudson, 1981), and that model poroelastic squirt flow (d,e,f) Chapman, 2003). Also shown is Δt^* predicted solely by velocity anisotropy effects, computed using an elastic tensor for Olivine (Abramson et al., 1997) and $Q = 50$. Results for other isotropic values of Q are shown in Supplementary Figure 1. Velocity and attenuation for P (blue), S1 (orange) and S2 (green) are calculated by solving the Christoffel equation (see text for details) assuming a $50\,\mathrm{km}$ thick medium and a dominant frequency of $0.1\,\mathrm{Hz}$. For the fluid-filled models we use an isotropic solid with velocities $v_P = 6.5\,\mathrm{km\,s^{-1}}$ and $v_S = 3.6\,\mathrm{km\,s^{-1}}$ which contains melt inclusions with $v_S = 2.7\,\mathrm{km\,s^{-1}}$, $\rho = 2700\,\mathrm{kg\,m^{-3}}$. These parameters are chosen to be broadly consistent with previous effective medium modelling of melt-induced seismic anisotropy (Hammond et al., 2014).

the medium. This broadly represents teleseismic shear-waves propagating through the upper mantle. As Δt^* represents the difference in attenuation between the fast and slow shear-waves there is a discontinuity at $\theta = 60°$ in the scattering model where the polarisation direction of S1 and S2 swap (Figure 1b,c). The importance of this is that crack scattering only predicts $\Delta t^* > 0$. It is also worth noting that the scattering model predicts non-physical negative 1/Q values for propagation angles around $\theta = 45°$ due to approximations used to calculate the imaginary components of the elastic tenor. This result can also be seen in Crampin (1984), where the approximations are developed.

The complex elastic tensor for the squirt flow model is calculated following the method of Chapman (2003). As a numerical example, we calculate velocity (Figure 1d), attenuation (Figure 1e), and Δt^* (Figure 1f) as a function of propagation angle for a frequency of $0.1\,\mathrm{Hz}$ using the same isotropic solid and crack fill properties and aspect ratio as before. Additionally, we specify a total porosity, $\Phi = 0.05$; a grain-sized microcrack density, $\epsilon_c = 0.05$; a meso-scale (i.e., larger than grain size) fracture density, $\epsilon_f = 0.1$; a fracture length, $a_f = 10\,\mathrm{m}$; and an aspect ratio, $r = 1 \times 10^{-4}$. Fracture and microcrack density are related to the respective porosities (or volume fractions) Φ_f and Φ_c in the squirt flow model by

$$\Phi_f = \frac{4}{3}\pi\epsilon_f r \tag{8}$$

and

$$\Phi_c = \frac{4}{3}\pi\epsilon_c r \tag{9}$$

(Chapman, 2003). This yields a fracture porosity $\Phi_f = 4.2 \times 10^{-5}$ and a microcrack porosity $\Phi_c = 2.1 \times 10^{-5}$, with the remaining porosity modelled as spherical pore spaces. An important assumption of the squirt flow model is that microcracks and pores interact with only one meso-scale fracture, which in turn requires a low fracture density to be valid. We use a mineral-scale relaxation time $\tau_m = 2 \times 10^{-5}\,\mathrm{s}$ and grain size $\zeta = 120 \times 10^{-6}\,\mathrm{m}$, which are taken from Chapman (2003)'s numerical example.

From these numerical examples, we can see that the inclusion of poroelastic squirt flow effects has a significant effect on the predicted seismic velocities and attenuation. Furthermore, squirt flow is sensitive to fracture length, with only a small range of fracture lengths producing measurable Δt^* for a given frequency (Figure 2). This frequency range is determined by the characteristic fracture relaxation frequency ω_f which is related to fracture length a_f by

$$\omega_f = \frac{\zeta}{a_f}\omega_m\,, \tag{10}$$

where

$$\omega_m = \frac{2\pi}{\tau_m}\,. \tag{11}$$

It follows that different frequencies will induce squirt flow in different fracture sizes (Supplementary Figure 2). In practice, the fractures will not have a uniform length and there will be a range of frequencies. In

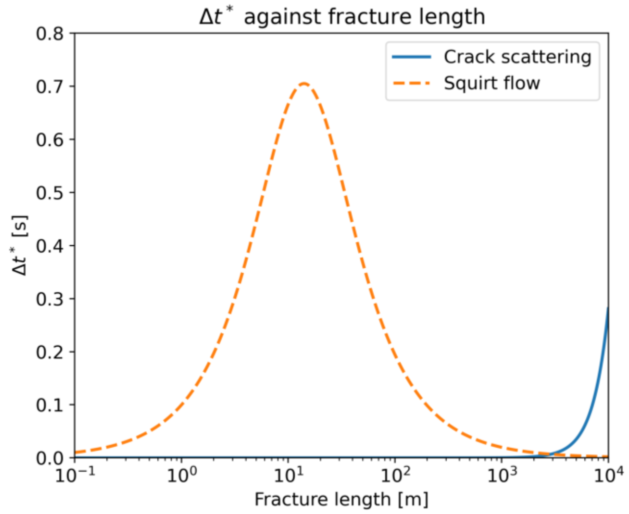


Figure 2 Anisotropic attenuation, Δt^*, as a function of fracture length as predicted by both squirt flow (dashed line Chapman, 2003) and crack scattering (solid line Hudson, 1980) models. Δt^* is calculated for a propagation angle $\theta = 70°$ relative to the crack normal and a frequency of $0.1\,\mathrm{Hz}$.

this modelling the frequency used ($0.1\,\mathrm{Hz}$) is assumed to be the dominant frequency of the seismic phases. The squirt flow model also assumes that the fractures are perfectly aligned. One effect of this assumption of identically sized and perfectly aligned fractures is that squirt flow predicts no attenuation anisotropy when $\theta = 90°$ (i.e., propagating parallel to the aligned fractures), whilst crack scattering predicts the maximum Δt^*. The squirt flow model produces a characteristic change in the polarisation of S1 and S2: in the example shown here this occurs at $\theta = 45°$, but the exact angle where this occurs depends on the model parameters used. Unlike the crack scattering mechanism, this allows for both positive and negative Δt^*. Observing this change of sign in Δt^* (or consistently observing $\Delta t^* < 0$) is a clear indicator of squirt flow and, therefore, of the presence of aligned fluid-filled fractures. This has been previously observed in microseismic datasets (Carter and Kendall, 2006; Usher et al., 2017). In particular, squirt flow could reasonably explain the results of Carter and Kendall (2006), who observed some cases where $\Delta t^* < 0$ in microseismic data recorded at the Valhall Field, in the Norwegian sector of the North Sea. Fractures on the order of $0.6\,\mathrm{m} - 6\,\mathrm{m}$ would produce attenuation anisotropy for microseismic frequencies (Supplementary Figure 2). Due to the length scales of both squirt flow and crack scattering, we would not expect significant attenuation anisotropy to occur for crystal lattice-preferred orientation mechanisms. The effects of velocity anisotropy, where S2 is more attenuated due to its larger travel time, are negligible (Supplementary Figure 1) and even if there are grain-scale fluid inclusions, such as grain boundary wetting, the squirt flow effects would occur well outside of the seismic frequency band. This, combined with the sensitivity of attenuation anisotropy to very low volume fractions of aligned fluid inclusions, makes measuring attenuation anisotropy a promising tool to detect fluids in the subsurface.

3 Instantaneous frequency as a measure of attenuation anisotropy

3.1 Instantaneous frequency

As we have shown, the crack scattering and squirt flow mechanisms both predict attenuation anisotropy which we could potentially measure in shear-wave splitting datasets. If the shear-waves S1 and S2 share the same source, geometrical spreading, and effective receiver transfer functions, then they should have equivalent frequency spectra if the intrinsic attenuation along the ray path is isotropic, barring the small difference caused by velocity anisotropy. Therefore, if we can measure a significant difference between the frequency content of each shear-wave, this might be attributed to attenuation anisotropy.

To measure the difference in attenuation between fast and slow shear-waves we apply the instantaneous frequency matching method of Matheney and Nowack (1995). Instantaneous frequency matching has been shown to be less sensitive to noise when measuring attenuation than taking spectral ratios (Matheney and Nowack, 1995; Engelhard, 1996). Instantaneous frequency matching also gives more robust estimates of isotropic mantle attenuation for teleseismic shear-wave phases than spectral ratios (Ford et al., 2012; Durand et al., 2013). This method also does not require the assumption of frequency-independent attenuation, which is useful for the case of fluid-filled fractures where frequency-dependent anisotropic attenuation is predicted even for seismic frequencies (e.g., Chapman, 2003; Jin et al., 2018). A similar approach can be taken by performing frequency shifts in the frequency domain (Quan and Harris, 1997), although we prefer the instantaneous frequency method as all measurements are kept in the time domain. Instantaneous frequency is a concept that arises from complex trace analysis (Gabor, 1946). A time-domain signal $x(t)$, such as a seismic wavelet, can be described in terms of its instantaneous amplitude (or envelope), $a(t)$, and instantaneous phase, $\theta(t)$,

$$x(t) = a(t)\cos\theta(t)\,, \tag{12}$$

which is equivalent to representing the signal by its complex Fourier spectrum (Engelhard, 1996). To construct the complex trace we apply a Hilbert transform to $x(t)$ to give the orthogonal quadrature (or imaginary) trace

$$y(t) = a(t)\sin\theta(t)\,, \tag{13}$$

with the complex trace then given by:

$$z(t) = x(t) + iy(t) = a(t)e^{i\theta(t)}\,. \tag{14}$$

From this complex trace, we then obtain the following expressions for instantaneous amplitude,

$$a(t) = [x(t)^2 + y(t)^2]^{(1/2)}\,, \tag{15}$$

and instantaneous phase

$$\theta(t) = \tan^{-1}\left(\frac{y(t)}{x(t)}\right). \tag{16}$$

The instantaneous frequency of our signal $x(t)$ is given by the rate of change of the instantaneous phase with respect to time

$$f(t) = \frac{1}{2\pi}\frac{d}{dt}\theta(t) \tag{17}$$

(Taner et al., 1979). This requires taking the derivative of an arctangent function, which results in

$$f(t) = \frac{1}{2\pi}\frac{x(t)\frac{d}{dt}y(t) - y(t)\frac{d}{dt}x(t)}{a(t)^2 + \epsilon^2}\,, \tag{18}$$

where ϵ is a damping factor that can be added to reduce the large positive and negative amplitude spikes that can occur (Matheney and Nowack, 1995). As we did not observe large spikes in our instantaneous frequency traces, and are only interested in a single time window, we did not add a damping factor. The instantaneous frequency values are also weighted by the squared instantaneous amplitude. This gives a damped and weighted instantaneous frequency within a specified analysis window as

$$f(t) = \frac{\int_{t-T}^{t+T} f(t')a(t')^2}{\int_{t-T}^{t+T} a(t')^2}. \tag{19}$$

When weighted by instantaneous amplitude the instantaneous frequency of a signal approaches the centre frequency, or spectral mean, of the signal's Fourier power spectra for a sufficiently large analysis window (Saha, 1987; Barnes, 1993). We use analysis windows picked for shear-wave splitting analysis, which isolate the phase of interest.

3.2 Instantaneous frequency matching of split shear-waves

The attenuation of a seismic phase is measured by matching the instantaneous frequency of the observed phase, f_{obs}, to that of a reference phase, f_{ref}. This is done by applying a frequency domain causal attenuation operator,

$$D(\omega) = \exp\left\{-\frac{\omega}{2}t^*\right\}\exp\left\{\frac{i\omega}{\pi}t^*\ln\frac{\omega}{\omega_r}\right\}, \tag{20}$$

where t^* is the anelastic delay time (2) and ω_r is the angular reference frequency (Muller, 1984), to the reference phase. Note that $D(\omega)$ affects both the amplitude and phase of the waveform, which has important effects when we relate attenuation anisotropy to shear-wave splitting. It is also worth noting that this causal attenuation operator is different from the operator stated in Matheney and Nowack (1995),

$$D(\omega) = \exp\left\{-\frac{\omega}{2}t^*\right\}\exp\left\{-\frac{i\omega}{\pi}t^*\ln\frac{\omega}{\omega_r}\right\}, \tag{21}$$

by a factor of -1 for the complex exponent term. Both expressions are valid, with the sign of the phase delay term being chosen to ensure that $D(\omega)$ produces a causal signal (Supplementary Figure 3). This depends on the choice of reference frequency and the sign convention of the fast fourier transform (FFT) implementation used. We choose to follow previous work (Ford

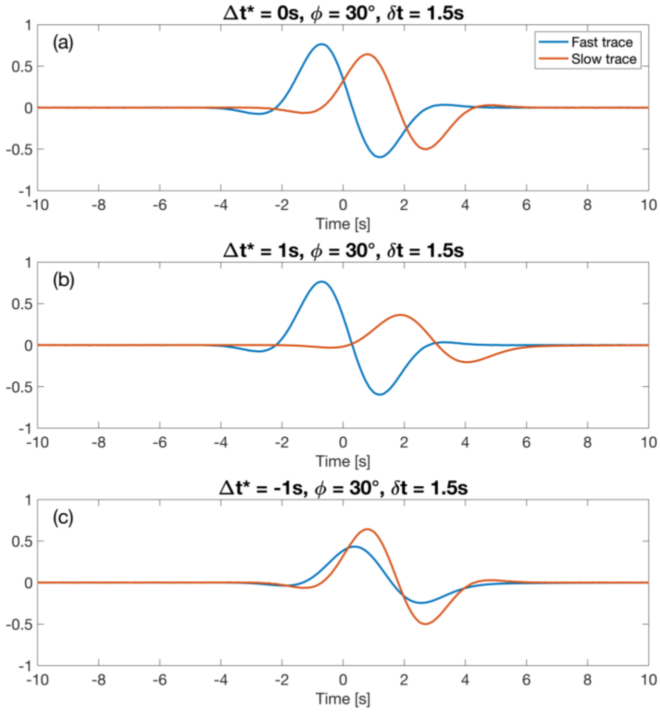


Figure 3 Example of Gabor wavelet synthetics used. (a) shows a synthetic where shear-wave splitting, with fast direction $\phi = 30^\circ$ and delay time $\delta t = 1.5\,\mathrm{s}$. (b) shows the synthetics in panel (a) where differential attenuation $\Delta t^* = 1.0\,\mathrm{s}$ has been applied by applying a causal attenuation operator to the slow shear-wave. (c) shows the synthetic from panel (a) where a differential attenuation $\Delta t^* = -1.0\,\mathrm{s}$ has been applied by attenuating the fast shear-wave. All synthetics in this Figure are generated with a source polarisation of 70°, a dominant frequency of $0.2\,\mathrm{Hz}$ and a sample rate of $50\,\mathrm{ms}$.

et al., 2012; Durand et al., 2013) in using (20). Following Muller (1984) the reference frequency is set to the Nyquist frequency, as this ensures $\omega < \omega_r$ and imposes a negative phase shift for all frequencies when using (20). Another common choice of reference frequency is 1 Hz (e.g., Futterman, 1962; Aki and Richards, 1980; Ford et al., 2012; Durand et al., 2013), which works provided that it is outside the frequency range of interest. One final important point to note, which may be slightly obfuscated by our choice of notation, is that this choice of attenuation operator implicitly assumes that Q is constant with frequency. This is a reasonably safe, and common, assumption to make for the seismic frequency band (e.g., Aki and Richards, 1980). However, this does mean that whilst there is no assumption of constant Q in the measurement of instantaneous frequency (Dasios et al., 2001; Ford et al., 2012), the common choice of $D(\omega)$ adds this assumption to the instantaneous frequency matching process.

Where a match in the instantaneous frequencies is achieved (i.e., $\Delta f = f_{ref} - f_{obs} = 0$) the t^* operator that is retrieved represents a differential attenuation between f_{ref} and f_{obs}. The physical meaning of the measured differential attenuation depends on the selection of f_{obs} and f_{ref}. For example, to measure lowermost mantle attenuation, the lower mantle transiting S phase can be used as a reference phase for ScS. The differential attenuation between the S and ScS phases can then be attributed to the divergence of the phases' ray paths in the lower mantle (Ford et al., 2012; Durand et al., 2013).

To measure attenuation anisotropy, instead of choosing a separate seismic phase as the reference phase we take advantage of shear-wave splitting and use one of the split shear waves as the reference phase. This gives the differential attenuation between S1 and S2, which we have previously described as Δt^* (equation 3). The sign of Δt^* indicates whether the fast (S1) or slow (S2) shear-wave has experienced more attenuation.

For this method to work, the fast polarisation direction must be correctly identified so that the fast and slow shear-waves can be separated. This is important as shear-wave splitting delay times are typically much smaller than the dominant period of the signal. This assumption is often made for teleseismic shear-waves (e.g., Silver and Chan, 1988; Chevrot, 2000). A consequence of this is that fast and slow shear-waves are not wholly split in time. This causes interference between the two shear-waves if they are viewed in the incorrect reference frame, which consequently affects the apparent frequency content of the two shear-waves. This makes the frequency content of each component dependent on the orientation of the reference frame. This is then further complicated by the phase shift introduced by the causal attenuation operator $D(\omega)$. We will expand on this further below, using example synthetic shear-waves.

3.3 Frequency domain effects of shear-wave component rotation and attenuation anisotropy

We can use synthetic data to explore the effects of component rotation and attenuation anisotropy on the frequency content of the apparent S1 and S2 phases, without the constraints attached to real observations of shear-wave splitting. All our synthetic examples are generated using a Gabor wavelet

$$x(t) = \cos(2\pi f_0(t - t_0) + \nu)\exp\{-4\pi^2 f_0^2 (t - t_0)^2/\gamma^2\}, \quad (22)$$

with a dominant, or carrier, frequency $f_0 = 0.2\,\mathrm{Hz}$ and a time shift $t_0 = 0\,\mathrm{s}$. The parameters γ and ν control the shape of the wavelet. For small γ the wavelet has a delta-like impulse and for large γ it has an oscillatory character. The parameter ν describes the symmetry of the wavelet. For $\nu = 0$, the wavelet is symmetric and when $\nu = \frac{-\pi}{2}$ or $\frac{\pi}{2}$ it is antisymmetric (Červený et al., 1977). Here we follow Matheney and Nowack (1995) and use the parameters $\gamma = 4.5$ and $\nu = 2\pi/5$. Synthetics are generated at a sample frequency of 20 Hz. The wavelet is then projected onto horizontal component seismograms from a desired initial source polarisation. Shear-wave splitting is applied to each synthetic by specifying the two desired shear-wave splitting parameters: the fast direction, ϕ_f, and delay time, δt. Where attenuation anisotropy is applied the synthetic is rotated to the fast polarisation direction, to isolate the fast and slow shear-waves, and either the slow trace or the fast trace is attenuated to achieve a positive or negative Δt^*.

Using simple synthetic examples (Figure 3), the effect that attenuation anisotropy has on shear-wave splitting can be seen. Here we generate synthetics with a fast polarisation direction of $30°$, a lag time of $1.5\,\mathrm{s}$ and a source polarisation of $70°$ (Figure 3a). The slow trace is attenuated by applying a causal attenuation operator (20) where $t^* = 1.0\,\mathrm{s}$, introducing a differential attenuation (or attenuation anisotropy), $\Delta t^* = 1.0\,\mathrm{s}$ (Figure 3b). A negative differential attenuation $\Delta t^* = -1.0\,\mathrm{s}$ can be introduced by instead attenuating the fast shear-wave (Figure 3c). Visual inspection of these synthetics shows a loss of amplitude on the attenuated trace. However, we can also observe an additional time delay in the attenuated traces introduced by the phase terms of the causal attenuation operator (equation 20), which has significant implications for measurements of shear-wave splitting.

The effect of component rotation on shear-wave frequency content can be further demonstrated using the synthetics from Figure 3a and 3b. The seismograms are rotated to the geographic reference frame (i.e., where a reference frame rotation $\phi_r = 0°$ returns the North and East components) and then rotated through reference frame angles in the range of $-90 \leq \phi_r \leq 90$. At each ϕ_r the amplitude of the frequency spectra is calculated, along with the instantaneous frequency within a $10\,\mathrm{s}$ analysis window centred on the wavelets (Figure 4). In the case where $\Delta t^* = 0\,\mathrm{s}$ the spectral amplitude of the fast (Figure 4a) and slow (Figure 4c) shear-waves vary with ϕ_r. When the fast and slow shear-waves are correctly separated at $\phi_r = 30°$ or $\phi_r = -60°$ there is no difference in the respective instantaneous frequencies, which are both measured as $0.2\,\mathrm{Hz}$. When $\Delta t^* = 1\,\mathrm{s}$ is applied the frequency content of the fast shear-wave should be unchanged, which is the case at $\phi_r = 30°$. The additional attenuation applied to the slow shear-wave reduces the effect of component rotation, but the effect is still strong enough to affect our instantaneous frequency matching method. These examples also show that instantaneous frequency retrieves the average amplitude-weighted frequency for each trace.

The synthetic shear-waves shown in Figure 3b, where $\phi_f = 30, \delta t = 1.5$, and $\Delta t^* = 1$, can also be used to demonstrate how instantaneous frequency matching can retrieve the applied attenuation anisotropy. Again the synthetic is initially rotated to the geographic reference frame and then rotated over the range $-90 \leq \phi_r \leq 90$. This simulates searching over the full range of reference frame rotations to test all potential fast shear-wave polarisations. At each reference frame rotation the instantaneous frequency of the two horizontal components is measured (Figure 5a), along with the difference in instantaneous frequencies (Figure 5b). The second eigenvalue of the trace covariance matrix, λ_2, is also calculated after correcting for the lag time $\delta t = 1.5\,\mathrm{s}$ (Figure 5c). We calculate λ_2 as it is commonly used in shear-wave splitting analysis that employs eigenvalue minimisation (Silver and Chan, 1991; Wuestefeld et al., 2010; Walsh et al., 2013). Shear-wave splitting introduces a phase delay between the two orthogonally polarised shear-waves, resulting in an elliptical particle motion. Eigenvalue minimisation methods use λ_2 to characterise this, or the seismic energy on the transverse component of a seismogram viewed in the radial-transverse reference frame. It therefore follows that if attenuation anisotropy introduces an additional phase delay term, this can add a source of systematic error to shear-wave splitting measurements. Only when the data is corrected for the applied Δt^* by attenuating the apparent fast shear-wave, which is the reference phase for a positive Δt^*, is the input fast polarisation direction able to be retrieved (Figure 5c). In this example, we know $\Delta t^* = 1\,\mathrm{s}$ and can omit a search over a range of potential Δt^* values.

The instantaneous frequency of the apparent fast and slow shear-waves varies as a function of reference frame rotation ϕ_r (Figure 4, 5a). For both the uncorrected (solid lines) and corrected (dashed) traces there are two points where the instantaneous frequencies match, which can be seen as minima in $|\Delta f|$ (Figure 5b). In the uncorrected data, these points are separated by approximately $90°$ and if $\Delta t^* = 0$ then one minima lies at the fast polarisation direction. When there is attenuation anisotropy these minima are not located at the true fast polarisation direction (solid line, Figure 5b). When the correction for Δt^* is applied these minima collapse towards one another, but do not necessarily converge to the same point.

Figure 5c shows the effect that attenuation anisotropy has on shear-wave splitting measurements, as characterised by λ_2. If we do not correct for the applied attenuation anisotropy then the λ_2 minima can appear to be less pronounced and deflected from the true fast polarisation direction. In this example, this synthetic shear-wave splitting has no clear λ_2 minimum when we correct for the imposed delay time of $1\,\mathrm{s}$. The minimum λ_2 occurs at a fast polarisation direction of $-78.24°$ compared to the true fast polarisation direction of $30°$. When we correct for Δt^* this effect is entirely removed and we can retrieve the input shear-wave splitting parameters. This error in fast polarisation direction increases with Δt^* and may not be fully captured by standard methods of measurement uncertainty estimation such as, for example, using the F-test derived 95% confidence region of the measured λ_2 values (Silver and Chan, 1991; Walsh et al., 2013), as the frequency effects of attenuation anisotropy distort λ_2 with rotation angle (Figure 5c). The magnitude of this effect depends on the strength of attenuation anisotropy.

3.4 Grid searching over component rotation and attenuation anisotropy to match instantaneous frequency

These synthetic examples (Figure 5) highlight an important challenge in measuring attenuation anisotropy for shear-waves. The inherent rotational interference between the fast and slow shear-waves makes measuring Δt^* highly dependent on accurately identifying the correct fast polarisation direction. Meanwhile, the error that Δt^* can introduce into shear-wave splitting measurements means that we cannot treat measurements of

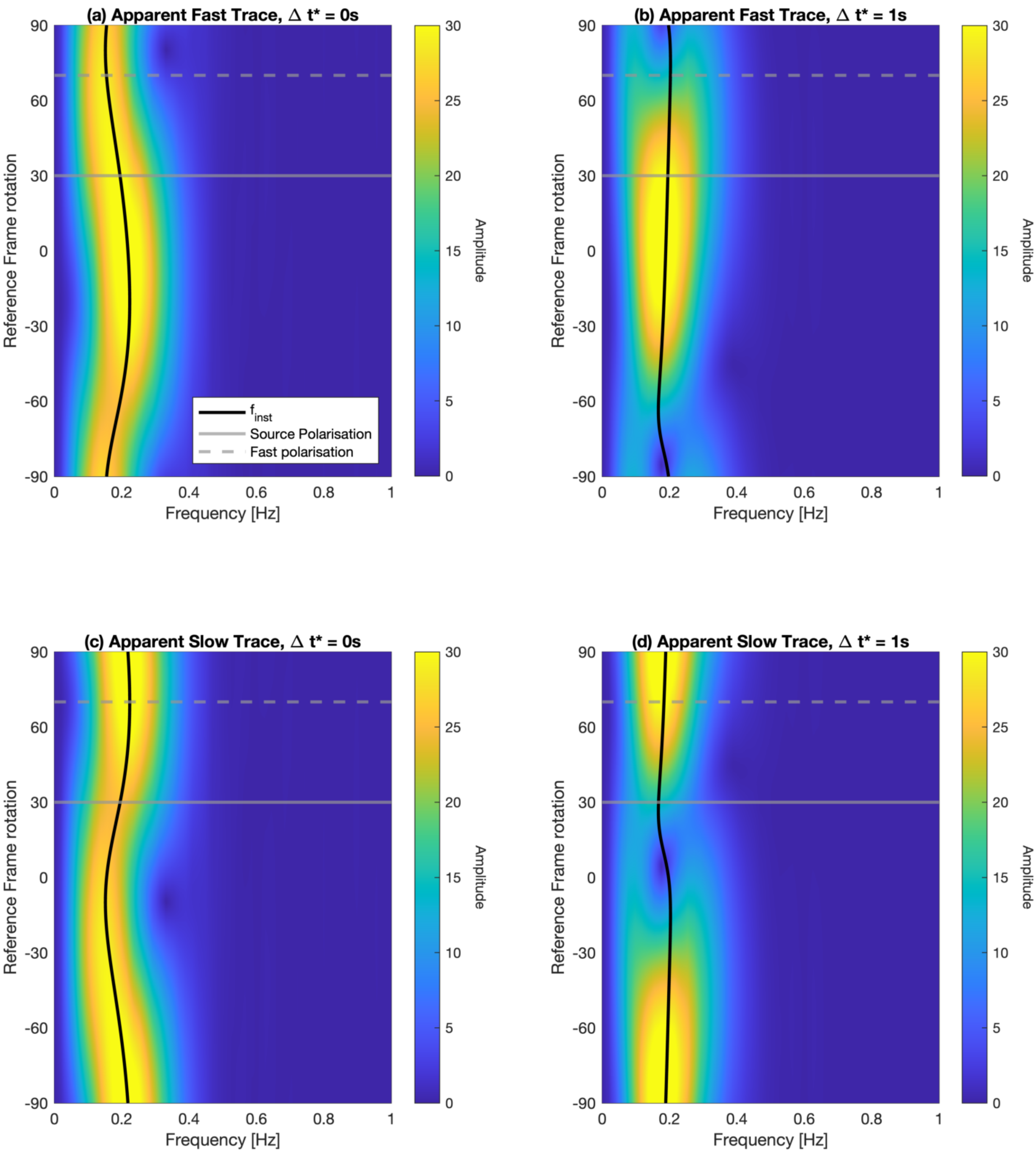


Figure 4 Amplitude spectra of synthetic shear-waves as a function of component reference frame rotation. At each reference frame rotation angle, we calculate the amplitude spectra and instantaneous frequency (black line) for the apparent fast and slow shear-waves. The left column shows the frequency content for the synthetics shown in Figure 3a, where $\phi_f = 30°, \delta t = 1.5\,\mathrm{s}$ and $\Delta t^* = 0\,\mathrm{s}$. The right column shows the frequency content for the synthetics shown in 3b, where an attenuation anisotropy of $\Delta t^* = 1\,\mathrm{s}$ has been applied to the synthetics shown on the right. This Δt^* is applied before rotating the components.

the fast polarisation direction as independent. To successfully measure Δt^* we must, therefore, also identify the true fast polarisation direction.

One strategy to achieve this is to search over both the potential component (or reference frame) rotation angles ϕ_r and differential attenuation Δt^*. To transform the instantaneous frequency matching process into a minimisation, simplifying the grid search, we adjust the objective function from Δf, used in Matheney and Nowack (1995) to $|\Delta f| = |f_{ref} - f_{obs}|$. In this form we have to fix the reference and observed traces to allow for automation of the grid search and set the apparent

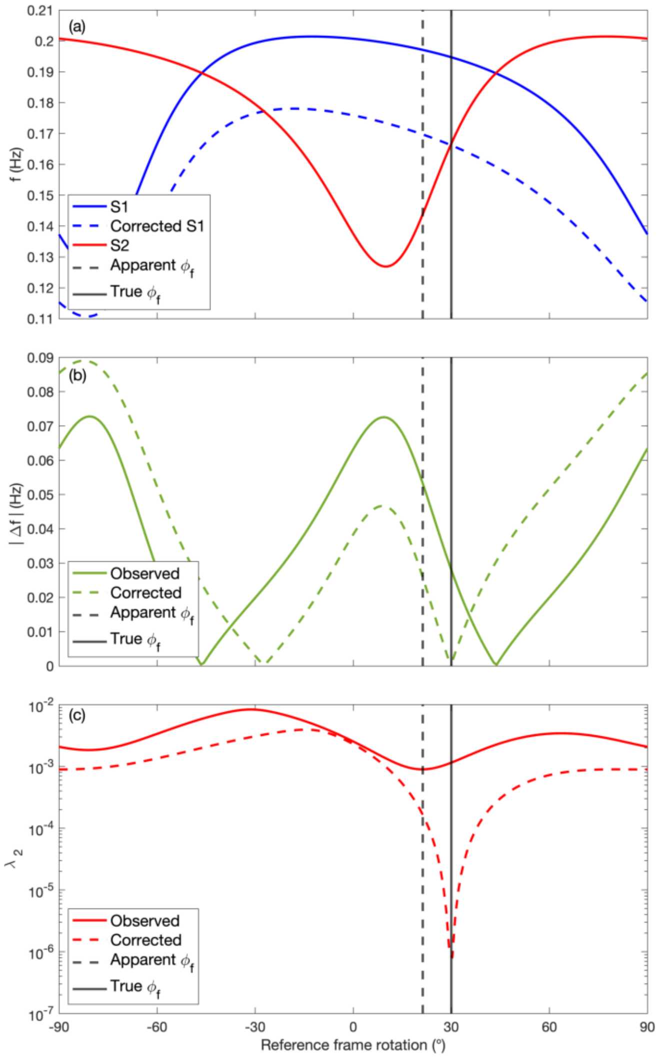


Figure 5 Example of the effects of component rotation and attenuation anisotropy on the frequency content and shear-wave splitting (parameterised by λ_2 for a synthetic shear-wave. Instantaneous frequency (a), the difference in instantaneous frequency (b), and the second eigenvalue of the trace covariance matrix (c) measured for the synthetic shear-wave shown in Figure 3b over the range of reference frame rotations of the horizontal components (solid lines). At each reference frame rotation, we then correct for the differential attenuation by attenuating the apparent fast shear-wave (blue) by $t^* = 1\,\mathrm{s}$ and repeat the measurements (dashed lines). The solid vertical line shows the applied (or true) fast polarisation direction, $30°$ and the dashed vertical line shows the fast polarisation direction that would be recovered if the synthetics are not corrected for Δt^*.

S1 phase as f_{ref} and the apparent S2 phase as f_{obs}, assuming that ϕ_r is the fast polarisation direction. This assumption means that we are unable to immediately determine the sign of Δt^* as cases where $\Delta t^* < 0$ are reported at the $90°$ from the true fast polarisation direction (i.e., the traces have been rotated such that S2 has become the reference phase). To find the correct sign for a Δt^* measurement we must return to input data, correct for Δt^* and then measure shear-wave splitting. If the measured fast polarisation agrees with ϕ_r, within measurement uncertainty, this indicates a positive Δt^*. If the difference between the fast polarisation and ϕ_r is approximately $90°$, within measurement uncertainty, this indicates that ϕ_r is the polarisation direction of the slow shear-wave which requires a negative Δt^*.

If we look at grid search results for individual shear-waves (Figure 6a), it becomes clear that we cannot uniquely constrain ϕ_r and Δt^* for a single event using our grid search method. One property of the relationship between the instantaneous frequency of split-shear waves and component rotation that we can take advantage of to resolve this is that instantaneous frequency (as a function of component rotation) is also dependent on the source polarisation of the shear-waves. Performing a grid search over ϕ_r and Δt^* for synthetics with example source polarisations of $45°$ (Figure 6a), $130°$ (Figure 6b) and $285°$ (Figure 6c), we can see that whilst we are unable to retrieve the input parameters $\phi_r = 30°, \Delta t^* = 1\,\mathrm{s}$ in each case there is a different subset of the model space which minimises $|\Delta f|$. For each source polarisation, this subset includes the true model parameters. When the examples are summed, the model space which can minimise $|\Delta f|$ is greatly reduced (Figure 6). In this simple, low noise example the minima of the sum returns the input $\phi_r, \Delta t^*$ exactly.

Therefore, we can measure ϕ_r and Δt^* if we have sufficient measurements of shear-waves with different source polarisations, where the assumption that all shear-waves sample the same attenuation anisotropy can be made. For this stacking method to work well, data with a good spread of source polarisations is desirable. For real data this does place constraints on where measurements can be made, as measuring shear-wave splitting from sources with an even distribution of source polarisations that sample a single region of attenuation anisotropy could be challenging.

4 Synthetic examples

We demonstrate our $|\Delta f|$ stacking method using synthetic shear-wave data. These examples show that our method can retrieve input shear-wave splitting and attenuation anisotropy parameters. As before, we use a Gabor wavelet and generate a set of 100 synthetic shear-waves. These synthetics are generated with a random source polarisation drawn from a continuous uniform distribution between $0°$ and $360°$ and with a dominant frequency drawn from $f \sim \mathcal{N}(0.1, 0.02)$. Shear-wave splitting, with a fast direction $\phi_f = 30°$ and delay time $\delta t = 1.0\,\mathrm{s}$, is applied to all synthetics. Attenuation anisotropy, with $\Delta t^* = 1\,\mathrm{s}$, is applied by attenuating the slow shear-wave. Random white noise with a noise fraction, or noise-to-signal ratio, of 0.075 is also added to the synthetics after rotating the components to the geographic reference frame. This represents a good signal-to-noise ratio, of ca. $13 : 1$ for real data as this example is intended to represent the ideal case for attenuation anisotropy measurements. To mimic the preprocessing of real data the synthetics are band-pass filtered, using a two-pole two-pass Butterworth filter with corners of $0.01\,\mathrm{Hz}$ and $0.3\,\mathrm{Hz}$. The absolute difference in instantaneous frequency, $|\Delta f|$, is calculated for candidate ϕ_r values over the range $-90° \leq \phi_r \leq$

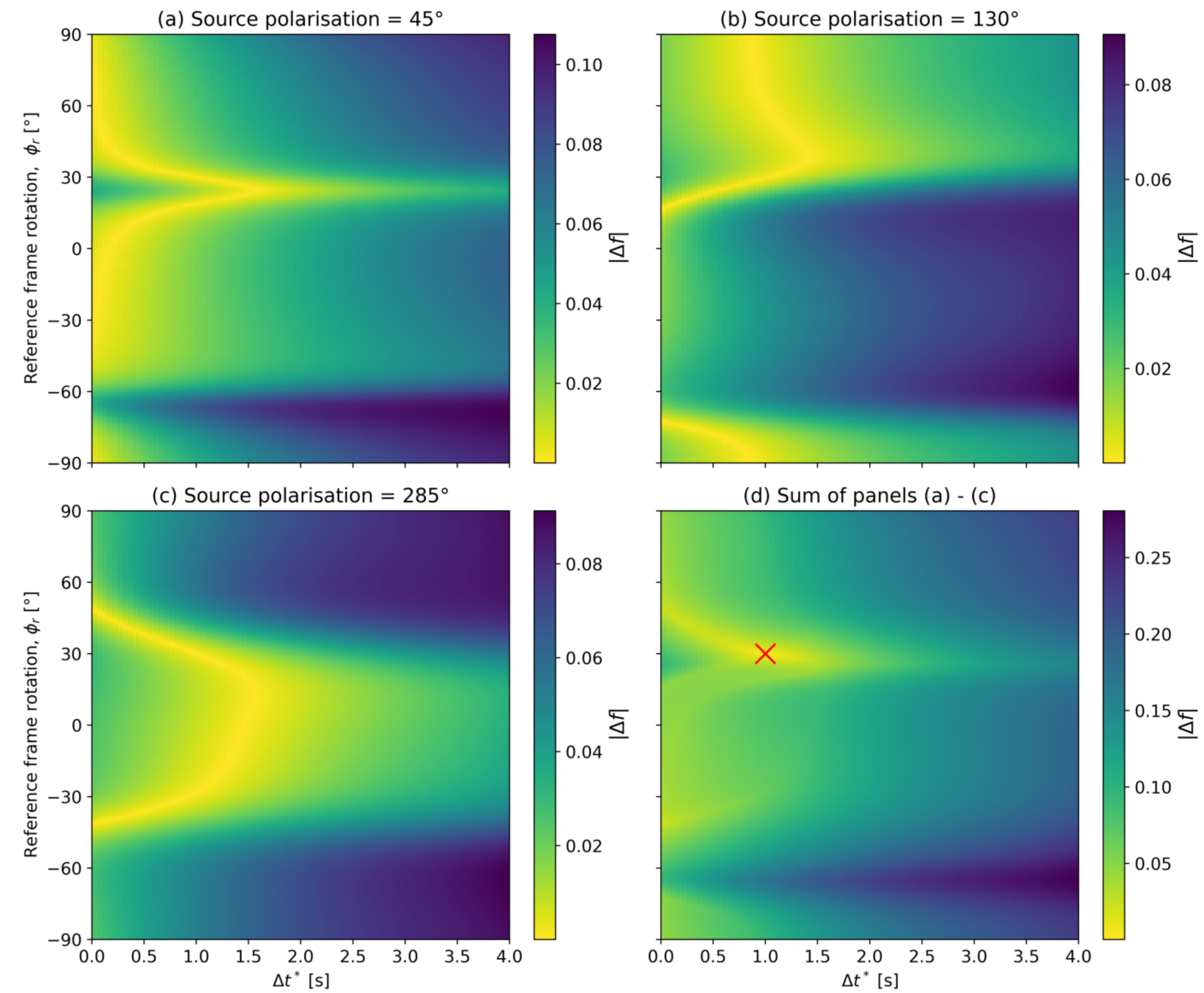


Figure 6 Example $|\Delta f|$ grid search results for individual synthetic waveforms generated at different source polarisations. Synthetics are generated with shear wave splitting parameters $\phi_f = 30°, \delta t = 1.0\,\mathrm{s}$, attenuation anisotropy $\Delta t^* = 1\,\mathrm{s}$ and source polarisations of $45°$ (a), $130°$ (b) and $285°$ (c). These $|\Delta f|$ surfaces can then be stacked (d), with the minima of the stack returning the input attenuation anisotropy Δt^* and fast polarisation direction ϕ_f.

$90°$, and candidate Δt^* in the range $0 \leq \Delta t^* \leq 4\,\mathrm{s}$ as shown in Figure 6. To account for potentially uneven source polarisation coverage, where data from one source polarisation could dominate the stack, we perform a weighted stacking similar to what can be used for shear-wave splitting (Restivo and Helffrich, 1999). Each $|\Delta f|$ grid is weighted by $1/N$, where N is the number of waveforms recorded in a $10°$ source polarisation bin. The best-fitting ϕ_r and Δt^* is found by taking the minima of the weighted stack (Figure 7).

To estimate the uncertainties in our measurements, we bootstrap our $|\Delta f|$ stacking. The 100 $|\Delta f|$ grids are bootstrap sampled, with replacement, 10,000 times. We repeat the source polarisation weighted stacking for each set of bootstrap samples. The resulting distribution of the minimum $|\Delta f|$ for each bootstrap sample (Figure 8) can be used to define a 95% confidence region in the stacked $|\Delta f|$. An upper-tailed test, where any $|\Delta f|$ that is below the 95% confidence threshold estimated from the bootstrapping (Figure 8) is considered to reasonably explain our data, is used. This 95% confidence threshold can then be mapped back onto weighted $|\Delta f|$ stack and estimate the uncertainties of $\phi_r, \Delta t^*$ from the length and width of the confidence region (Figure 7), following a similar approach to shear-wave splitting studies (e.g., Wuestefeld et al., 2010; Walsh et al., 2013; Hudson et al., 2023). If the minimum of the weighted $|\Delta f|$ stack sits outside of this confidence threshold, then this tells us that there is either data polluting the stacks that require removal, or that we are unable to confidently measure Δt^* for that station.

For the synthetic examples attenuation anisotropy parameters $\phi_r = 31 \pm 1°, \Delta t^* = 0.95 \pm 0.16°$ are measured for the synthetics where $\Delta t^* = 1\,\mathrm{s}$ was imposed (Figure 7a). In the case where $\Delta t^* = -1\,\mathrm{s}$ was added, we instead measure $\phi_r = -56 \pm 1°$ and $\Delta t^* = 1.00 \pm 0.08\,\mathrm{s}$ (Figure 7b). These results show that the source polarisation stacking method can correctly, and accurately, measure the attenuation anisotropy parameters

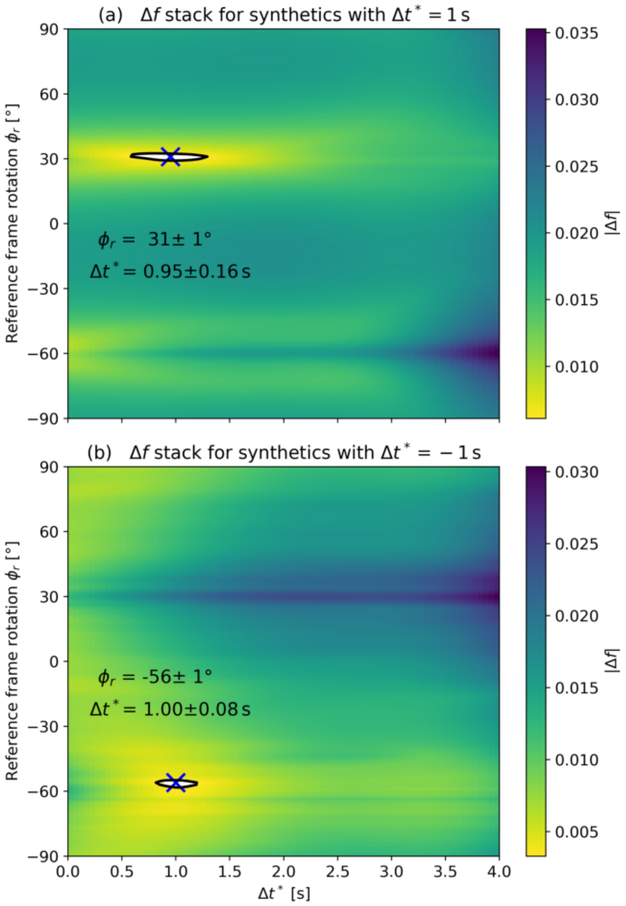


Figure 7 Source polarisation weighted, stacked $|\Delta f|$ surfaces. Each panel shows the $|\Delta f|$ stack measured for 100 Gabor wavelet synthetics generated with shear-wave splitting parameters $\phi_f = 30°, \delta t = 1.0\,\mathrm{s}$ and $\Delta t^* = 1\,\mathrm{s}$ (a) or $\Delta t^* = -1\,\mathrm{s}$ (b). $|\Delta f|$ is calculated for each synthetic by grid searching over ϕ_r and Δt^*. The delay time δt is not measured at this point in the workflow as it does not affect $|\Delta f|$ provided that a suitable analysis window has been chosen. Each synthetic is generated with a random source polarisation and with a dominant frequency drawn from $f \sim \mathcal{N}(0.1, 0.02)$.

$\phi_r, \Delta t^*$. It is worth noting that we are not able to exactly retrieve the input parameters as we are only correcting for the difference in frequency content between the fast and slow shear-waves and are not removing the effect of attenuation, which results in a permanent loss of amplitudes. The negative Δt^* example (Figure 7b) shows the expected result from imposing $\Delta t^* > 0$ in the grid search. The change in sign is instead mapped into the reference frame rotation, with the minimum $|\Delta f|$ being approximately $90°$ rotated from the fast polarisation direction. This has the effect of setting the slow shear-wave as the assumed reference (less attenuated) phase and the fast-shear wave as the observed (more attenuated) phase. This allows for the measurement of both positive and negative Δt^*, which is important to enable us to distinguish between potential mechanisms of attenuation anisotropy.

In these synthetic examples, the sign of Δt^* is known. For real data and experiments, we do not necessarily have this *a priori* information. Determining the sign of Δt^* is very important to measuring attenuation anisotropy as it allows us to distinguish between crack scattering and squirt flow mechanisms (Figure 1c,f). Observing negative Δt^* is potentially a powerful diagnostic for the presence of subsurface fluids, as it cannot be explained by velocity anisotropy and requires attenuation anisotropy due to a more complex mechanism such as squirt flow. In turn, squirt flow requires very small volume fractions of fluids hosted by aligned fractures to generate a measurable Δt^*. To correctly find the sign of Δt^* the most convenient approach is to measure attenuation anisotropy (ϕ_r and Δt^*) and then use these results to remove the effect of attenuation anisotropy before measuring shear-wave splitting. The measured shear-wave splitting parameters, after correcting for attenuation anisotropy, will tell us the correct fast polarisation direction. If the measured fast polarisation agrees with ϕ_r, within measurement uncertainty, this indicates a positive Δt^*. If the difference between the fast polarisation and ϕ_r is approximately $90°$, within measurement uncertainty, this indicates that ϕ_r is the polarisation direction of the slow shear-wave which requires a negative Δt^*.

This can be demonstrated by measuring shear-wave splitting for two synthetic datasets, where the positive Δt^* synthetics are generated with $\phi_f = 30°, \delta t = 1.0\,\mathrm{s}, \Delta t^* = 1\,\mathrm{s}$ and the negative Δt^* synthetics are generated using $\phi_f = 30°, \delta t = 1.0\,\mathrm{s}, \Delta t^* = -1\,\mathrm{s}$. Here shear-wave splitting is measured before (Figure 9a,c) and after (Figure 9b,d) correcting for the previously measured attenuation anisotropy (Figure 7). Shear-wave splitting is measured using eigenvalue minimisation as implemented in the analysis code SHEBA (Wuestefeld et al., 2010). The individual shear-wave splitting results are then stacked, with each result weighted by the signal-to-noise ratio and the number of measurements within a $10°$ back azimuth bin (Restivo and Helffrich, 1999).

The results of our shear-wave splitting measurements highlight two key factors. Firstly, the subtle effects that attenuation anisotropy has on apparent shear-wave splitting are clear. In the case with a positive Δt^*, where the slow shear-wave is more attenuated, the additional phase shift caused by the attenuation anisotropy nearly doubles the delay time relative to the true value (Figure 9a,b). The opposite occurs for a negative Δt^*. When the fast shear-wave is more attenuated it is delayed by the phase term of the attenuation operator, which reduces the delay time. In this example, the effect is sufficiently strong to delay the ‘fast’ shear-wave such that it arrives after the ‘slow’ shear-wave, which causes the $90°$ rotation in the apparent fast polarisation direction (Figure 9c). In both cases, after correcting for the measured attenuation anisotropy (Figure 7) we can retrieve the input shear-wave splitting parameters with significantly higher accuracy than if no correction had been applied (Figure 9b,d).

In the previous example the input splitting parameters are constant across the synthetics, but it is common to get some scatter in individual shear-wave splitting measurements. To test the effect of this on our source

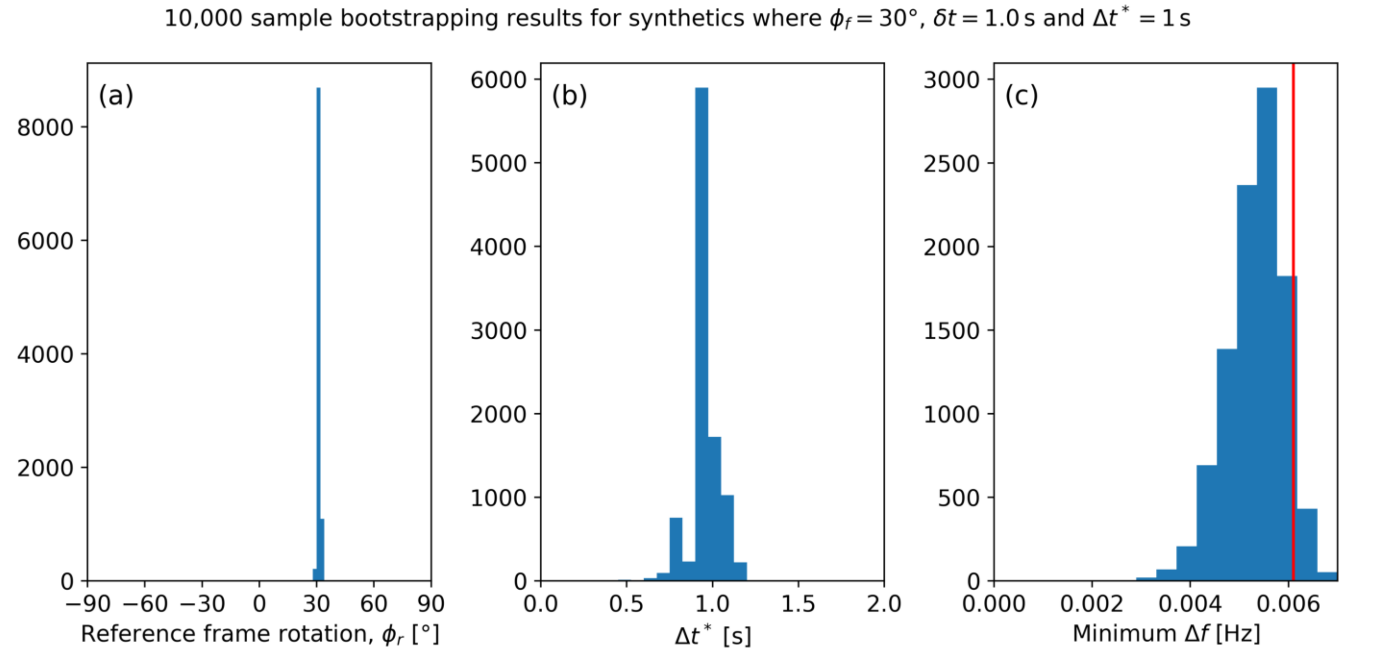


Figure 8 Bootstrapped summary statistics for the parameters ϕ_r, (a) and Δt^* (b) along with the minimum $|\Delta f|$ of each bootstrapped stack (c) for synthetic $|\Delta f|$ stacking example shown in Figure 7a. The initial set of 100 individual $|\Delta f|$ measurement grids is resampled, with replacement, 10,000 times and we repeat the stacking for each sample. The red vertical line in panel (c) indicates the bootstrap estimated 95% confidence level in $|\Delta f|$.

polarisation stacking method we repeat the previous experiment and instead randomly draw 100 samples for ϕ_f and δt from $\phi_f \sim \mathcal{N}(30, 5)$ and $\delta t \sim \mathcal{N}(1.5, 0.15)$. This set of shear-wave splitting parameters are then used to generate two sets of synthetics as previously described, where we apply $\Delta t^* = 1\,\mathrm{s}$ to one set and $\Delta t^* = -1\,\mathrm{s}$ to the other. In these cases we are unable to perfectly retrieve the input attenuation anisotropy, measuring $\phi_r = 30 \pm 1°, \Delta t^* = 0.95 \pm 0.30°$ and $\phi_r = -62 \pm 1°, \Delta t^* = 0.90 \pm 0.08°$ (Supplementary Figure 4). However when we correct the synthetics using the best-fitting Δt^* and ϕ_r we still significantly improve the accuracy of the shear-wave splitting measurements (Supplementary Figure 5). Again this highlights the effect that attenuation anisotropy can have on shear-wave splitting measurements and that after correcting for attenuation anisotropy, even though the corrections are not perfect, we are broadly able to retrieve the true shear-wave splitting parameters (Supplementary Figure 5b,d) even with some scatter in the individual observations. This does, however, increase the uncertainty in the retrieved shear-wave splitting and it should be noted that in both cases it is not possible to exactly retrieve the input shear-wave splitting parameters.

5 Measuring shear-wave splitting and attenuation anisotropy for FURI, Ethiopia

To demonstrate the potential of Δt^* to detect melt or fluids in the subsurface we choose the station FURI, which is situated on the margin of the Main Ethiopian Rift (MER) close to Addis Ababa. FURI is operated as part of the Global Seismograph Network (Albuquerque Seismological Laboratory/USGS, 2014). We choose this locality as previous SKS shear-wave splitting studies have interpreted seismic anisotropy due to aligned melts beneath the MER (e.g., Ayele et al., 2004; Kendall et al., 2005; Bastow et al., 2010; Hammond et al., 2014). Melt has also been inferred by seismic tomography, using body waves (e.g., Bastow et al., 2008), Rayleigh waves (e.g., Chambers et al., 2022) and ambient noise (e.g., Chambers et al., 2019; Eshetu et al., 2021), receiver functions (e.g., Rychert et al., 2012) and magnetotelluric (Whaler and Hautot, 2006) studies. Aligned melt mechanisms should also produce a strong signal of attenuation anisotropy (Figure 1,2), making the MER and surrounding region a natural target to search for attenuation anisotropy. FURI is one of the few permanent stations in the region, with over 20 years of waveform data available, making it a good station to measure SKS shear-wave splitting. SKS is an ideal phase to attempt attenuation anisotropy measurements using our source polarisation stacking method. As SKS travels through the outer core as a P-wave, it is only sensitive to anisotropy after it exits the core. Arriving near vertically beneath the station, we can assume that all SKS phases sample the same region of the upper mantle regardless of backazimuth. Furthermore, SKS is radially polarised when exiting the core due to the P-to-S conversion (Hall et al., 2004). Therefore the backazimuthal coverage at FURI (Figure 10) approximately maps onto the achievable source polarisation coverage. Whilst this source polarisation coverage is not ideal (Supplementary Figure 6), it is sufficent to ensure that the $|\Delta f|$ stacking is stable. For other shear-wave phases, such as teleseismic or local S, it may not be possible to achieve sufficent source-polarisation

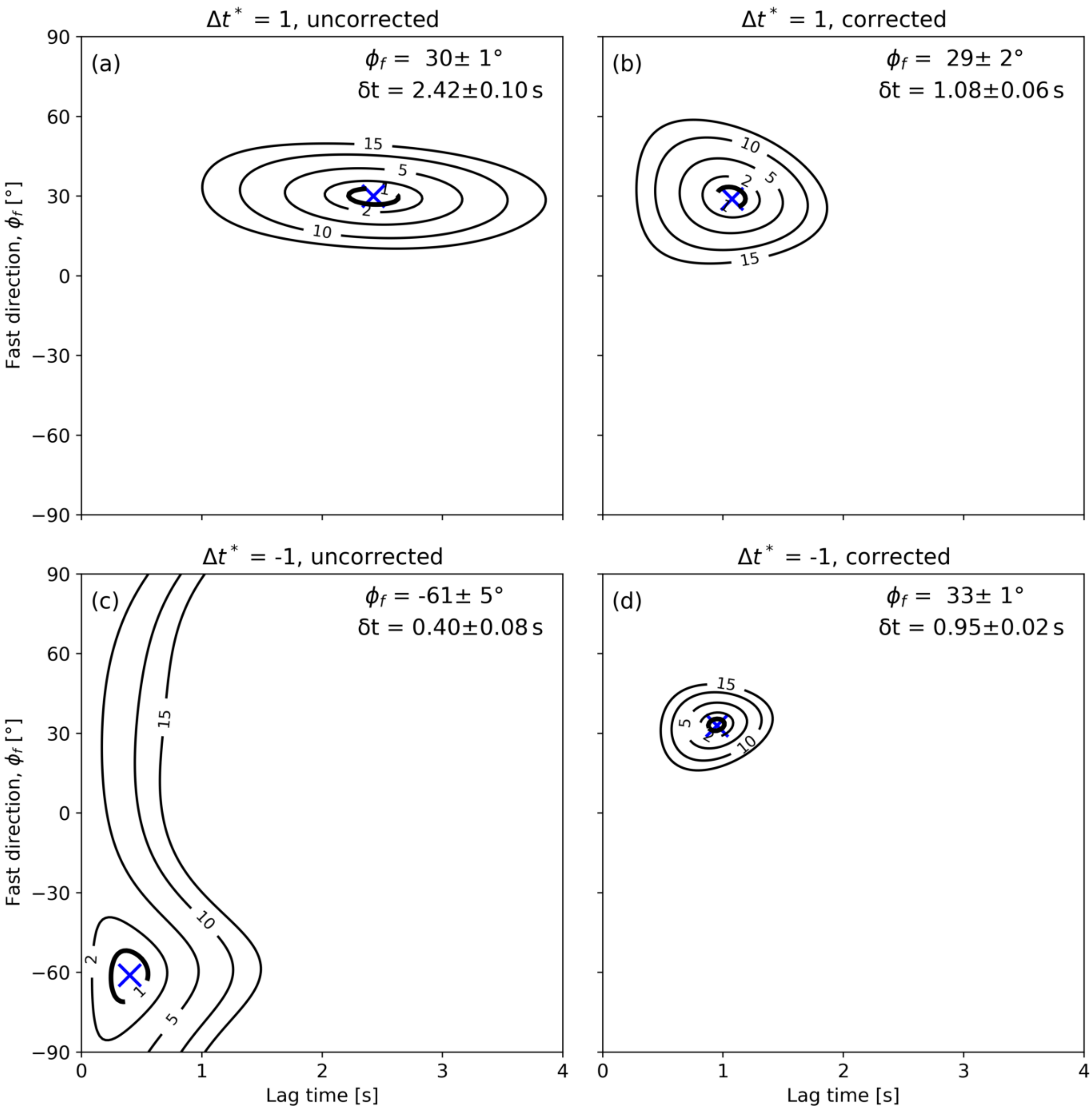


Figure 9 Results of synthetic shear-wave splitting measurement stacking, following the method of Restivo and Helffrich (1999). We generate 100 synthetics with shear-wave splitting parameters $\phi_f = 30^\circ$ and $\delta t = 1.5\,\mathrm{s}$. Attenuation anisotropy of $\Delta t^* = 1\,\mathrm{s}$ (a,b) or $\Delta t^* = -1\,\mathrm{s}$ (c,d) is applied. Panels (a,c) show the shear-wave splitting results if we do not correct for this attenuation anisotropy. Panels (c,d) show the result after we correct the synthetic data using measurements of $\phi_r, \Delta t^*$ made using our source polarisation stacking method. The stacked λ_2 surfaces are normalised by the 95% confidence value, indicated by the bold contours, which is derived from an F-test (Silver and Chan, 1991; Restivo and Helffrich, 1999).

coverage without relying on raypaths that sample different regions. In that case, our source polarisation stacking method cannot be applied unless it is clear the S phases are sampling the same region of attenuation anisotropy. Removing the requirement for source polarisation stacking is, therefore, desirable and is an avenue for future research.

It is worth noting that two layers of anisotropy have been suggested across the Main Ethiopian Rift, with the upper layer interpreted as aligned melt pockets and the lower layer associated with density-driven mantle flow due to the African superplume (Hammond et al., 2014). As only the upper layer is likely to host aligned melt inclusions, we do not expect the two-layer problem to have a significant effect on our results. However, it is worth noting that the contribution from the lower layer will introduce additional frequency mixing. We would not expect this to mask the strong attenuation anisotropy predicted for aligned melt inclusions, but this complication may increase the uncertainty in Δt^*.

Data is collected for 584 earthquakes, which are at a

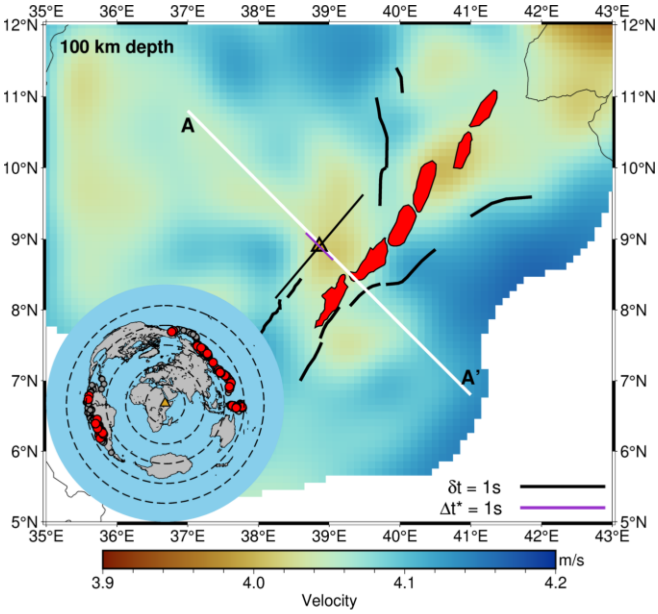


Figure 10 Map showing shear-wave velocity beneath the Main Ethiopia Rift at a depth of 100km, obtained from the joint inversion of ambient noise and teleseismic Rayleigh waves (Chambers et al., 2022). Thick black lines indicate border faults and red polygons indicate magmatic segments. The location of FURI is shown by the yellow triangle. Station averaged SKS shear-wave splitting, after correcting for attenuation anisotropy, indicated by the black bar plotted on FURI, where the length of the bar corresponds to the delay time, δt, and its orientation to the fast polarisation direction. Measured attenuation anisotropy is shown by the magenta bar and follows the same plotting convention as the shear-wave splitting result. The cross-section A-A' (white line) through the tomography model is shown in Figure 16. The inset map shows the locations of the 584 events used in this study (grey circles). From these events, we can identify 73 that yield clear SKS picks which are used to measure shear-wave splitting and attenuation anisotropy, shown by the red circles. We only consider events with an epicentral distance between $95°$ and $110°$, the dashed lines mark the distance from FURI (yellow triangle) in intervals of $30°$. Event locations are taken from the International Seismological Centre (2023) bulletin.

sufficient epicentral distance ($95°$ to $120°$) for SKS to be visible, recorded at FURI (Figure 10). Only earthquakes with a moment magnitude in the range of $5.5 \leq M_w \leq 7.0$ and a minimum depth of $50\,\mathrm{km}$ are used. All earthquake data were requested from the International Seismological Centre (2023) bulletin, with the dataset covering 21 years, from 1st January 2001 to 1st January 2022. Before analysis, all waveforms are corrected for instrument response and we detrend and demean the data using tools available in ObsPy (Beyreuther et al., 2010).

Shear-wave splitting is measured for all 584 SKS waveforms before measuring attenuation anisotropy. Whilst it is not essential to measure shear-wave splitting before attenuation anisotropy, and indeed we have already shown that attenuation anisotropy can affect shear-wave splitting measurements (Figure 4, 9), it can be a useful first step in analysis and enables us to manually inspect the waveforms data quality before measuring attenuation anisotropy. The waveform data is filtered using a two-pass two-pole Butterworth filter, with corner frequencies of $0.01\,\mathrm{Hz}$ and $0.3\,\mathrm{Hz}$. This enables a direct comparison of our results with previous SKS shear-wave splitting station averages (Ayele et al., 2004). The filtered waveforms are visually inspected and analysis window start/end search ranges are picked for waveforms where a clear SKS phase can be picked. This manual inspection reduces the dataset to 73 waveforms where SKS can be clearly identified. Figure 11a shows an example SKS phase used. We then measure shear-wave splitting using the shear-wave splitting analysis code SHEBA (Wuestefeld et al., 2010), which utilises the method of Silver and Chan (1991) as updated by Walsh et al. (2013). The optimum shear-wave splitting analysis window, which will also be utilised to measure attenuation anisotropy, is found using cluster analysis (Teanby et al., 2004). At this stage in shear-wave splitting analysis, one might seek to further reduce the dataset, by applying data quality thresholds based on Wuestefeld et al. (2010)'s shear-wave splitting quality parameter Q (which is not related to the attenuation quality factor) or by removing results which have large measurement errors in ϕ_f or δt (e.g., Kendall et al., 2005). In this case, we do not want to reduce the size of our dataset as this may remove data that exhibits attenuation anisotropy.

As in the synthetic shear-wave example, the station averaged shear-wave splitting is calculated by summing normalised second eigenvalue surfaces weighted by signal-to-noise ratio and source polarisation (Restivo and Helffrich, 1999). Our station averaged results of $\phi_f = 38 \pm 6°$ and $\delta t = 1.15 \pm 0.28\,\mathrm{s}$ are consistent, within uncertainty, with previously measured values of $\phi_f = 36 \pm 1°$ and $\delta t = 1.38 \pm 0.02\,\mathrm{s}$ (Figure 13a, Ayele et al., 2004).

For each SKS phase a $|\Delta f|$ surface is measured by grid searching over $-90° \leq \phi_r \leq 90°$ and $0\,\mathrm{s} \leq \Delta t^* \leq 4\,\mathrm{s}$, in intervals of $1°$ and 0.05 respectively. These measurements use the analysis windows previously defined, using Teanby et al. (2004)'s cluster analysis method, for the corresponding shear-wave splitting measurement. As outlined previously we stack our $|\Delta f|$ measurements, weighted by source polarisation. Measurement uncertainties are determined by bootstrapping the stacking process as described for the synthetic examples (Supplementary Figure 7). The source polarisation of each waveform is estimated in the shear-wave splitting measurement process by SHEBA (Wuestefeld et al., 2010). From the stacked $|\Delta f|$ the measured attenuation anisotropy is $|\Delta t^*| = 0.45 \pm 0.20\,\mathrm{s}$ and $\phi_r = -45 \pm 3\,\mathrm{s}$. As in the synthetic examples, we are not immediately able to determine the sign of Δt^*. To find the correct sign, each SKS phase must be corrected for the measured attenuation anisotropy. Then the shear-wave splitting of the corrected waveforms, with the effects of attenuation anisotropy removed, can be measured. The attenuation anisotropy corrections are applied by rotating the waveforms to ϕ_r and attenuating the retrieved reference phase by the measured $|\Delta t^*|$. An example corrected SKS phase is shown in Figure 11b. This has the effect of removing the phase shift introduced by at-

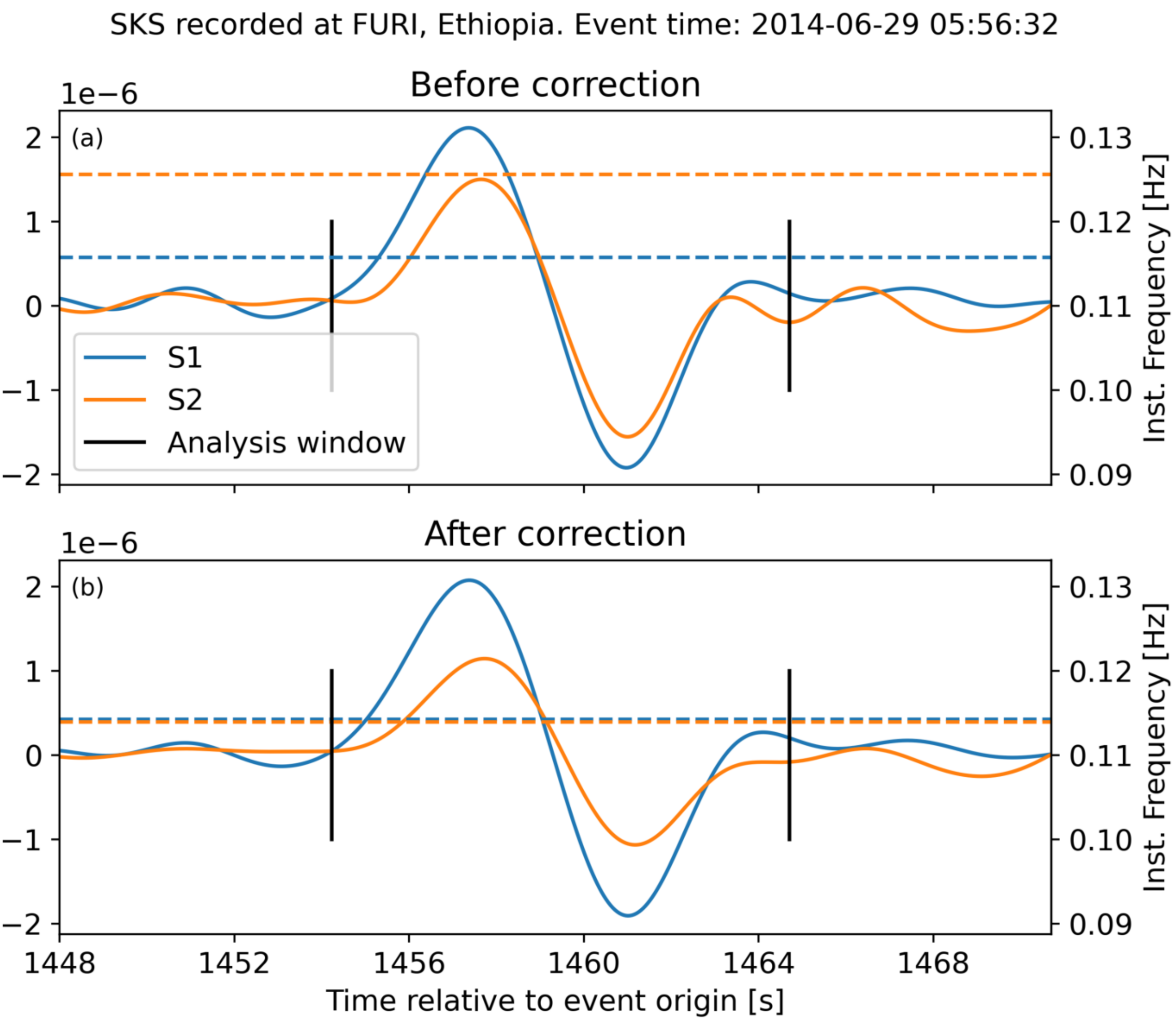


Figure 11 One of the SKS phases recorded at FURI, Ethiopia, which we use to measure attenuation anisotropy. Panel (a) shows the pre-processed SKS phase rotated to the station averaged fast polarisation direction, $\phi = 38°$, that we measure for FURI and time shifted by $\delta t = 1.15\,\mathrm{s}$. The fast shear-wave, S1, is shown in blue and the slow shear-wave, S2, is shown in orange. Note that S1 appears to have a slightly longer period than S2, which suggests it has been more attenuated. Dashed lines show the measured instantanous frequency for the chosen analysis window (black lines). Panel (b) is plotted in the same style, showing the SKS phase after we correct for the measured attenuation anisotropy.

tenuation anisotropy, although as this is a station averaged measurement the correction may not be perfect. There will be a permanent loss of amplitudes, but the difference in frequency content between the fast and slow shear-waves should be removed (Figure 11b) and this will not affect measurements of shear-wave splitting.

After correcting the SKS waveforms, we measure station averaged shear-wave splitting of $\phi_f = 40 \pm 5°$ and $\delta t = 1.60 \pm 0.34\,\mathrm{s}$. This result is also consistent with previous work, within measurement uncertainties, but the best-fitting delay time has increased by $0.45\,\mathrm{s}$. As the difference between ϕ_f and ϕ_r is $90°$, within measurement uncertainty, we interpret that the fast shear-wave has been more attenuated than the slow shear-wave and that $\Delta t^* < 0$. This gives a final joint measurement of station averaged shear-wave splitting and attenuation anisotropy at FURI of $\phi_f = 40 \pm 5°, \delta t = 1.60 \pm 0.34\,\mathrm{s}$ and $\Delta t^* = -0.45 \pm 0.20\,\mathrm{s}$.

6 Characterising fluid inclusions using velocity and attenuation anisotropy

In our examples, using both synthetic and real data, we have established that we can measure attenuation anisotropy in split shear-waves. Our observation of attenuation anisotropy in real SKS data for FURI, Ethiopia is an important result and corroborates previous work which has interpreted seismic anisotropy in terms of preferentially oriented melt inclusions both beneath FURI (Ayele et al., 2004) and potentially more broadly across the Main Ethiopian Rift (Kendall et al., 2005; Bastow et al., 2010). The additional measurement of attenuation anisotropy gives us further insight into this mech-

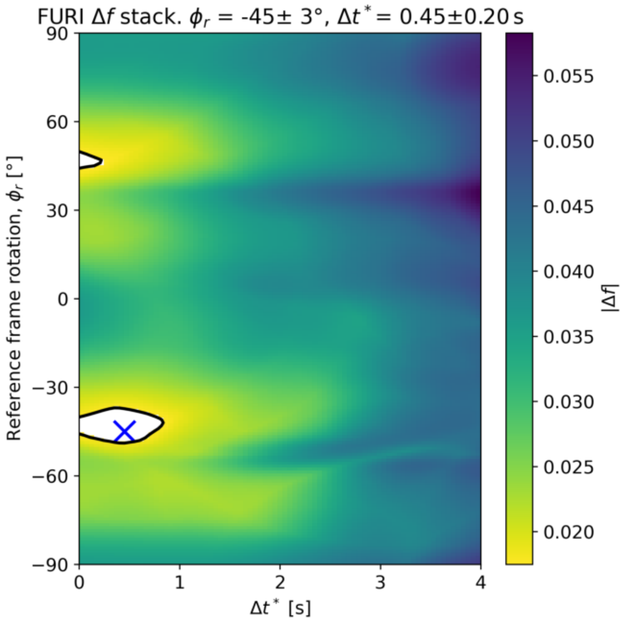


Figure 12 Source polarisation weighted, stacked $|\Delta f|$ surface for FURI, Ethiopia. This result is obtained by stacking 73 $|\Delta f|$ surfaces measured for SKS waveforms recorded at FURI. We measure an attenuation anisotropy of $\phi_r = -45 \pm 3°$ and $\Delta t^* = 0.45 \pm 0.20\,\mathrm{s}$, indicated by the blue cross. The 95% confidence region in our solution is demarcated by the bold contour and coloured white. Our approach to measuring $|\Delta f|$ means that we cannot initially determine the sign of Δt^*. Upon analysis of our corrected SKS shear-wave splitting results (Figure 13b) we can determine that $\Delta t^* = -0.45 \pm 0.20\,\mathrm{s}$.

Model parameter	Value
Melt fraction	0.01
Fracture density	0.1
Micro-crack density	0
Aspect ratio	1×10^{-4}
Solid P wave velocity, v_P	$6.2\,\mathrm{km\,s^{-1}}$
Solid S wave velocity, v_S	$3.6\,\mathrm{km\,s^{-1}}$
Solid density, ρ	$2700\,\mathrm{kg\,m^{-3}}$
Melt P wave velocity, v_P	$2.7\,\mathrm{km\,s^{-1}}$
Melt density, ρ	$2700\,\mathrm{kg\,m^{-3}}$

Table 1 Parameters used in squirt flow modelling of attenuation anisotropy observed at FURI, Ethiopia. For details of microscale relaxation time, τ_m, grain size, ζ, and fracture length, a_f, used see text.

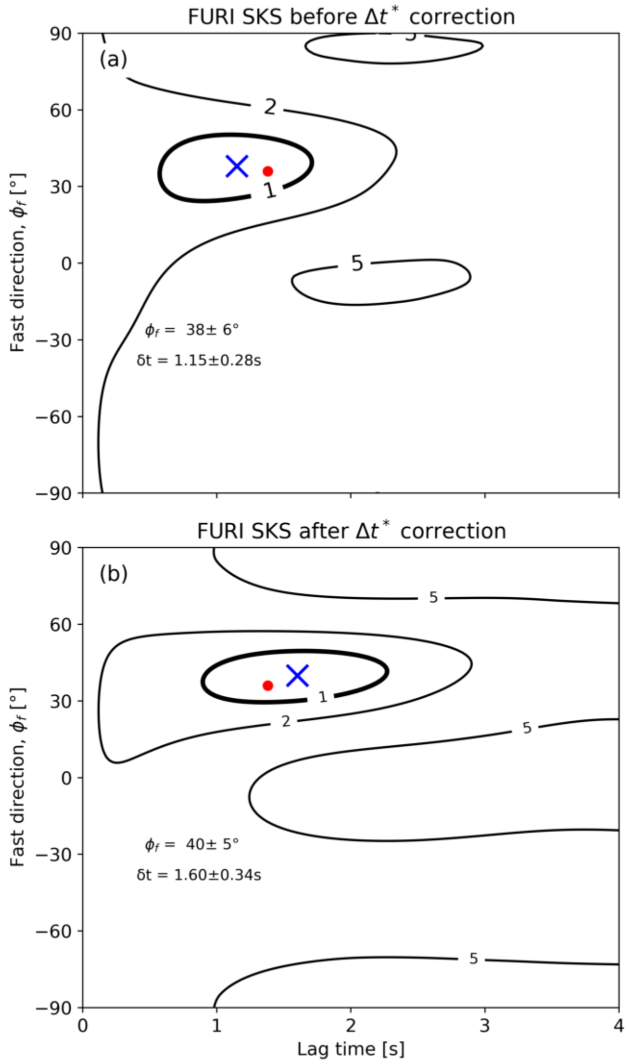


Figure 13 Station averaged shear-wave splitting for FURI, Ethiopia, plotted similarly to Figure 9. Shear-wave splitting measurements are stacked using the Restivo and Helffrich (1999) method before (a) and after (b) correcting for measured attenuation anisotropy of $\Delta t^* = 0.45\,\mathrm{s}$ and $\phi_r = -45°$. The red dot shows the previous station averaged SKS shear-wave splitting measurement ($\phi_f = 36 \pm 1°, \delta t = 1.38 \pm 0.03\,\mathrm{s}$) at FURI (Ayele et al., 2004, red circle).

anism. The observation of $\Delta t^* = -0.45\,\mathrm{s}$ can only be explained by the poroelastic squirt flow of a fluid-filled medium given that alternate mechanisms, such as crack scattering or velocity anisotropy effects, always predict that the slow shear-wave should be more strongly attenuated and $\Delta t^* > 0$ (Figure 1e,f). Furthermore, our modelling of attenuation anisotropy due to crystal preferred orientation of Olivine shows that, for reasonable mantle Q, the expected Δt^* is at least one order of magnitude smaller than what we observe (Figure 1c,f, Supplementary Figure 1). We would also expect other potential mechanisms for intrinsic attenuation anisotropy for crystal preferred orientation, such as grain boundary melt squirt, to operate at frequencies significantly above the seismic frequency band. This allows us to discount crystal preferred aligment mechanisms, making attenuation anisotropy a good indicator for the presence of aligned fluid inclusions.

With the observed $\Delta t^* = -0.45 \pm 0.20\,\mathrm{s}$ strongly suggesting the presence of aligned fluid inclusions, the natural next question is how can we characterise these inclusions. As we have already described, the squirt flow model requires a large set of parameters to characterise a fluid-filled fractured medium. One of the most important parameters to have reasonable constraints on is mineral relaxation time, τ_m, which is empirically derived and is proportional to the viscosity of the saturating fluid and inversely proportional to the permeability of the host rock (Chapman et al., 2003). Previous work inverting shear-wave splitting for fracture models using the squirt flow model has shown that the inversion is highly sensitive to the τ_m used (Al-Harrasi et al., 2011). It has also been shown that varying τ_m has a substantial effect on the expected frequency-dependent seismic velocity anisotropy (Baird et al., 2013). Greater constraints on plausible values for τ_m in the upper mantle are required to enable detailed modelling of fracture characteristics. Any modelling of fracture properties is also dependent on the choice of grain size, ζ, and fracture length, a_f. Together τ_m, ζ and a_f describe the fracture scale squirt-flow relaxation time,

$$\tau_f = \frac{a_f}{\zeta}\tau_m, \tag{23}$$

which is also expressed as the squirt-flow frequency $\omega_f = \frac{2\pi}{\tau_f}$ and determines the frequency range of the fracture-dependent squirt flow effects. Whilst this trade-off makes it difficult to constrain the fracture or grain size, if some reasonable assumptions are made it is still possible to constrain potential fracture orientations.

To search for potential fracture orientations, given the lack of constraint on τ_m we make some assumptions to simplify the problem. Outside of τ_m, fracture length, and grain size, there are 9 other potential free parameters required to calculate a complex elastic tensor using Chapman (2003)'s squirt flow model. We fix these parameters to the values in Table 1, which leaves fracture strike, dip, and medium thickness as free parameters to search over. Seismic velocities and densities are chosen to be consistent with previous effective medium modelling of the region (Hammond et al., 2014). A total porosity, or melt fraction, of 1%, is chosen, along with a fracture density of 0.1, as previous work suggests SKS shear-wave splitting at FURI could be explained by a melt fraction $\leq 1\%$ (Ayele et al., 2004). This represents a parsimonious choice of model parameters as we seek to explain our observations with a small melt fraction, where the implied fracture porosity (i.e., melt volume fraction hosted in the fractures) $\phi_f = 4.2 \times 10^{-5}$. If SKS is assumed to be vertically incident, then the fracture dip corresponds to the angle to fracture normal used in the earlier numerical examples (Figure 1), and the fracture strike is predominately controlled by the measured fast polarisation direction. This assumption also makes ray path length interchangeable with medium thickness.

We search for the best-fitting medium thickness, l, in the range $50\,\mathrm{km} \leq l \leq 150\,\mathrm{km}$ and fracture dip angle, θ, in the range $0^\circ \leq \theta \leq 90^\circ$ by rotating the elastic tensor to θ and calculating the predicted delay time, δt and attenuation anisotropy, Δt^*. The misfit for these predicted parameters is calculated using a normalised least-squares approach. To reflect the lack of constraint on τ_m, and therefore also τ_f, in upper mantle conditions this exercise is repeated over a large range of τ_m values, $10 \times 10^{-6}\,\mathrm{s} \leq \tau_m \leq 10 \times 10^{-2}\,\mathrm{s}$, an assumed grain size of $1\,\mathrm{mm}$ and fracture lengths of $10\,\mathrm{m}$, $100\,\mathrm{m}$ and $1000\,\mathrm{m}$. Figure 14 shows the τ_f required by the current choice of grain size, fracture length and τ_m (Figure 14a) along with the misfit of the best-fitting model (Figure 14b), predicted Δt^* and δt (Figure 14c,d), and the best-fitting medium thickness (Figure 14e) and fracture dip (Figure 14f) for a given τ_m. This modelling exercise can be repeated by fixing an assumed fracture length and varying the chosen grain size, which gives similar results (Supplementary Figure 8). Despite the lack of constraint on τ_m one set of model parameters emerges: a medium thickness in the range ca. $90\,\mathrm{km} - 120\,\mathrm{km}$ and fracture dip in the range ca. $38^\circ - 48^\circ$ which can reasonably explain the observed delay times and δt^*. It is worth noting that different modelled fracture lengths and grain sizes require a different range of τ_m values to fit the results. Therefore with better constraints on τ_m it would be possible to identify plausible fracture (or melt inclusion) lengths for a given grain size. This modelling also shows the value of measuring attenuation anisotropy. In addition to identifying the presence of aligned fluid-filled fractures, measurements of Δt^* add important constraints to fracture orientation. The uncertainty in the measurement of $\delta t = 1.60 \pm 0.34\,\mathrm{s}$ means that it can be reasonably explained by all τ_m (Figure 14d) and the additional measurement of Δt^* adds an extra data point. This uncertainty largely maps into melt volume fraction, which has a strong effect on the seismic velocity anisotropy, which we have elected to fix at $1\,\%$, and fracture density, which is required to be low and fixed to 0.1 The measured delay time can also be fitted by shallowly dipping or near-vertical fractures, with the addition of attenuation anisotropy, $\Delta t^* = -0.45 \pm 0.20\,\mathrm{s}$, requiring shallowly dipping fractures (Figure 1, Supplemental Figure 9). This relies on the assumption that the squirt flow model (Chapman, 2003) is valid in upper mantle conditions, where this model has not previously been tested. Poroelastic squirt flow requires that the melt is hosted in very low aspect ratio inclusions, due to the limitations of Eshelby's theory (Eshelby, 1957), and that the melt inclusions are near perflectly aligned. To our knowledge poroelastic squirt flow is the only model can explain the negative Δt^* observed. Furthermore, grain-scale melt squirt in the mantle has long been used to model isotropic velocity and attenuation (e.g., Mavko and Nur, 1975; Hammond and Humphreys, 2000). Adding melt squirt of aligned melt inclusions allows us to consider the contributions to velocity and attenuation anisotropy, but this requires that we can model low aspect ratio melt inclusions as fluid-filled fractures. Future work is needed to establish the theoretical attenuation anisotropy of larger aspect ratio melt inclusions, such as melt tubules.

The best-fitting fracture strike direction is found by setting the medium thickness to the thinnest plausible value from the previous modelling exercise, $90\,\mathrm{km}$, and searching over fracture dip and strike angles, where we seek to fit $\Delta t^*, \delta t$ and ϕ_f again using a normalised least-square cost function (Figure 15). This layer thickness is broadly consistent with previous estimates of the thickness of anisotropy beneath FURI (Ayele et al., 2004), al-

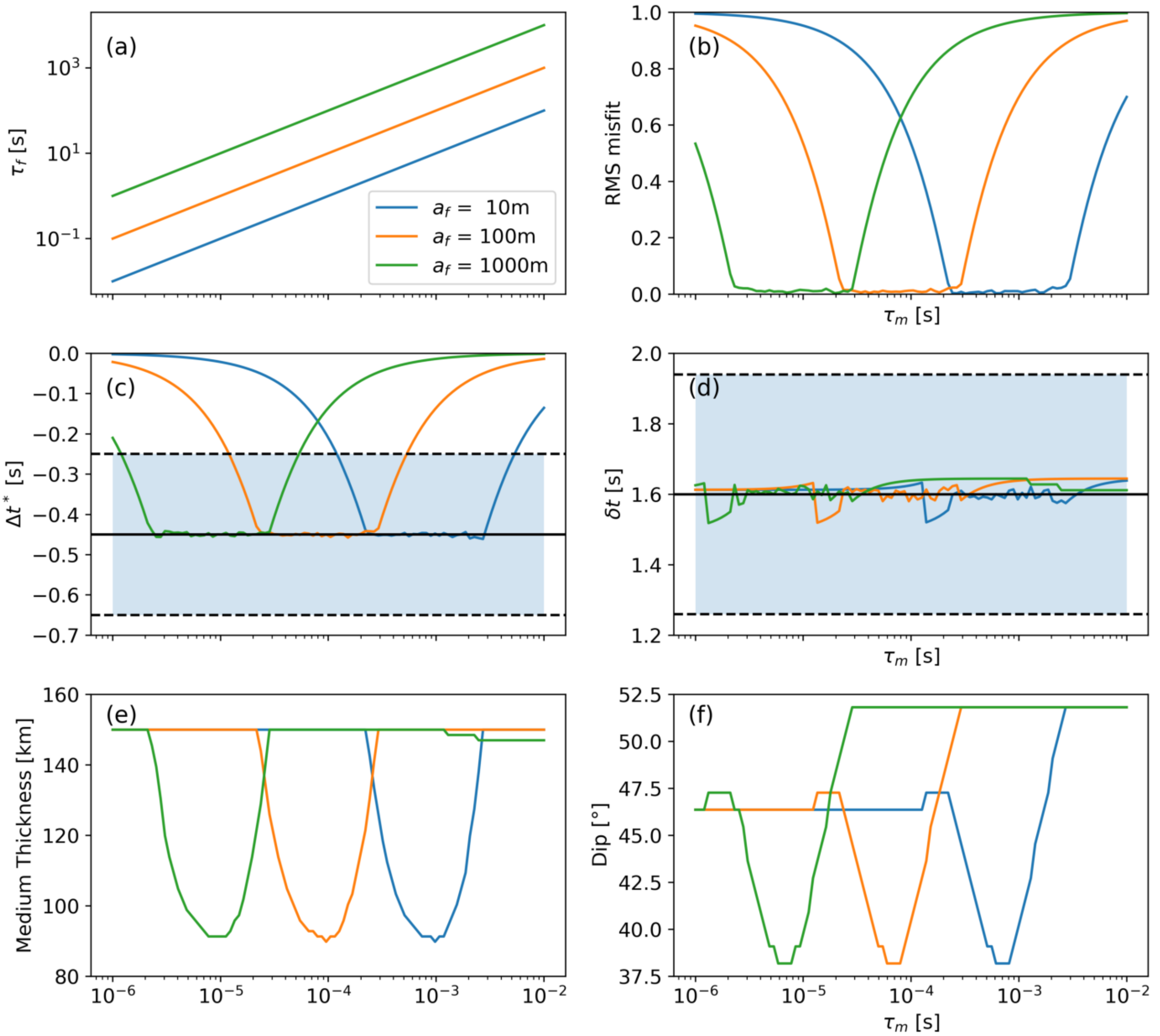


Figure 14 Results of modelling the fracture dip and medium thickness, or ray path length, which best explain the observed δt and Δt^* at FURI, Ethiopia using a squirt flow model (Chapman, 2003). Due to the lack of constraint on the mineral-scale relaxation time, τ_m, we search over a range of τ_m values for an assumed grain size of $1\,\mathrm{mm}$ and fracture lengths of $10\,\mathrm{m}$ (blue), $100\,\mathrm{m}$ (orange) and $1000\,\mathrm{m}$ (green). Panel (a) shows the fracture-scale relaxation time, τ_f, which is proportional to τ_m (23). The normalized least-square misfit of the best-fitting model for each τ_m is shown in (b), with the predicted Δt^* and δt shown in (c) and (d). The observed $\Delta t^* = -0.45\,\mathrm{s}$ and $\delta t = 1.6\,\mathrm{s}$ are shown by the solid black lines in (c) and (d), with the measurement uncertainties indicated by the shaded region. Panels (e) and (f) show the medium thickness, assuming a single anisotropic layer, and fracture dip angle required.

though a thinner region of melt inclusions could be accommodated by increasing the melt fraction. We assume a fracture length of $100\,\mathrm{m}$ and a grain size of $1\,\mathrm{mm}$ and set the mineral scale relaxation time, τ_m, to $9.55 \times 10^{-5}\,\mathrm{s}$. The best-fitting orientations give a fracture with a dip of $39°$ and an NW-SE strike (Figure 15). This rift perpendicular fracture orientation complicates previous interpretations that seismic anisotropy across the MER is due to rift parallel, vertical melt inclusions in the uppermost mantle (e.g., Ayele et al., 2004; Kendall et al., 2005). It is worth noting that it is only the addition of Δt^* which requires shallowly dipping fractures. This shallow dipping fracture model can then only fit the observed fast polarisation direction, $\phi_f = 40°$ if the fractures have a NW or SE strike. Alternatively, melt could be accomodated in inclusions with aspect ratios above the limit set by Eshelby (1957). This scenario is not modelled here, and would require revisiting long standing assumptions of aligned melt mechanisms for seismic anisotropy in the region as low aspect ratio melt inclusions have been the prevailing interpetation (e.g., Ayele et al., 2004; Kendall et al., 2005; Bastow et al., 2010; Hammond et al., 2014). A final possiblity is that there is some other, as yet unknown, mechanism for attenuation anisotropy in the uppermost mantle which can explain our observations whilst allowing for near-vertical

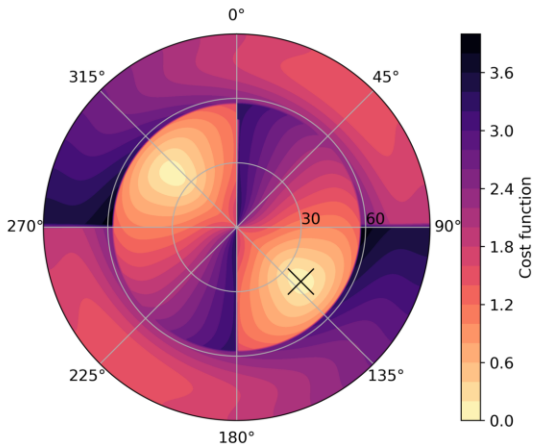


Figure 15 Forward modelling, assuming a poroelastic squirt-flow model (Chapman, 2003): results for fracture strike and dip that can explain the observed attenuation anisotropy and shear-wave splitting in vertically incident SKS phases at FURI. We model a medium with 1% total melt volume fraction, of which 0.1% is hosted in aligned fractures. Medium thickness is fixed to $90\,\mathrm{km}$ following our previous models (Figure 14e), which is consistent with previous estimates of the maximum thickness of anisotropy in the region (Ayele et al., 2004; Hammond et al., 2014). We use a normalised least squares cost function for Δt, δt, and ϕ_f. Our results favour shallowly dipping, 39°, fractures which are oriented NW-SE, which is approximately perpendicular to the Main Ethiopian Rift. For details of other model parameters used see text.

melt inclusions.

With only one data point we cannot say if this negative attenuation anisotropy, and the requirement for shallowly dipping fractures, is localised to FURI or is more widespread across the MER. Comparison to a recent shear-wave velocity tomography model at a depth of $100\,\mathrm{km}$ does indeed show a low-velocity anomaly which extends perpendicular to the rift axis directly beneath FURI (Figure 10; Chambers et al., 2022). A linearly interpolated cross-section through the model (Figure 16) shows that the feature is up to $70\,\mathrm{km}$ thick and situated directly beneath FURI. This feature could represent a network of shallowly dipping aligned melt inclusions extending away from the MER beneath FURI, which is causing the observed attenuation anisotropy. This is slightly thinner than what our models find, but this could potentially be accommodated by a modest increase in the overall melt fraction or fracture density. The melt fraction and fracture density were fixed to $1\,\%$ and 0.1 to simplify the modelling done here, but could plausibly be increased. Further work, such as siting several additional stations further along the anomaly perpendicular to the MER, is required to more thoroughly test if there are shallowly dipping melt inclusions extending away from the MER and to better constrain the extent of melt present. A current limitation is that most deployments at the MER are temporary and therefore often do not record a sufficiently large sample size of SKS phases for our stacking approach to be robust.

This example serves to highlight the potential of attenuation anisotropy to enhance our understanding of melt or fluid-rich regions, even where we have a good understanding of seismic anisotropy in the region. At a minimum attenuation anisotropy is potentially a useful tool for identifying the presence of fluids in the subsurface, even at very low volume fractions. More extensive, dense, measurements of shear-wave splitting and attenuation anisotropy may, in the future, allow for strong constraints to be placed on important properties such as the volume fraction of melt present and the orientation of the melt inclusions.

Currently our source polarisation stacking method can only be readily applied for SKS and other core-transiting shear-wave phases. In these cases we can make the assumption that all phases sample the sample region of the upper mantle beneath the station and, for some stations, achieved sufficient source polarisation coverage to measure attenuation anisotropy. For other shear-wave phases, such as local or teleseismic S, this is not the case. To achieve the requisite source polarisation coverage would require data that most likely samples different regions of anisotropy which makes taking a station average unsuitable. This poses a particular challenge to measuring shear-wave attenuation anisotropy in the near surface, where the potential for attenuation anisotropy to improve characterisations of fluid-filled fracture systems could prove powerful. Removing the requirement for source polarisation stacking is, therefore, desirable and is a promising avenue for future research.

7 Conclusion

Seismic attenuation anisotropy is a phenomenon which can be efficiently generated by models of fluid-filled fractures, particularly a squirt flow model. This attenuation anisotropy has a clear theoretical and observable effect on measurements of shear-wave splitting. The effect of attenuation anisotropy on the frequency content of split-shear waves can be measured using an adaptation of existing instantaneous frequency matching methods (Matheney and Nowack, 1995). Using synthetic shear-wave examples and SKS phases recorded at FURI, Ethiopia, we show these effects and that we can measure attenuation anisotropy and retrieve the underlying shear-wave splitting parameters. To explain the observed attenuation anisotropy, where the fast shear-waves appear more attenuated than the slow shear waves in SKS phases, a squirt flow model (Chapman, 2003) is required. Even allowing for a lack of constraints on the rock physics parameters it is clear that this requires shallowly dipping (ca. 40°) melt inclusions which strike perpendicular to the Main Ethiopia Rift. Whilst the modelled strike and dip of the melt inclusions is contrary to expectations from previous work (e.g., Ayele et al., 2004; Kendall et al., 2005; Bastow et al., 2010; Hammond et al., 2014), there is some potential correlation with low shear-wave velocity anomalies seen in recent tomographic models that extend away from the rift. These results highlight the power of attenuation anisotropy measurements as a blunt tool to detect the presence of aligned melt inclusions within the Earth.

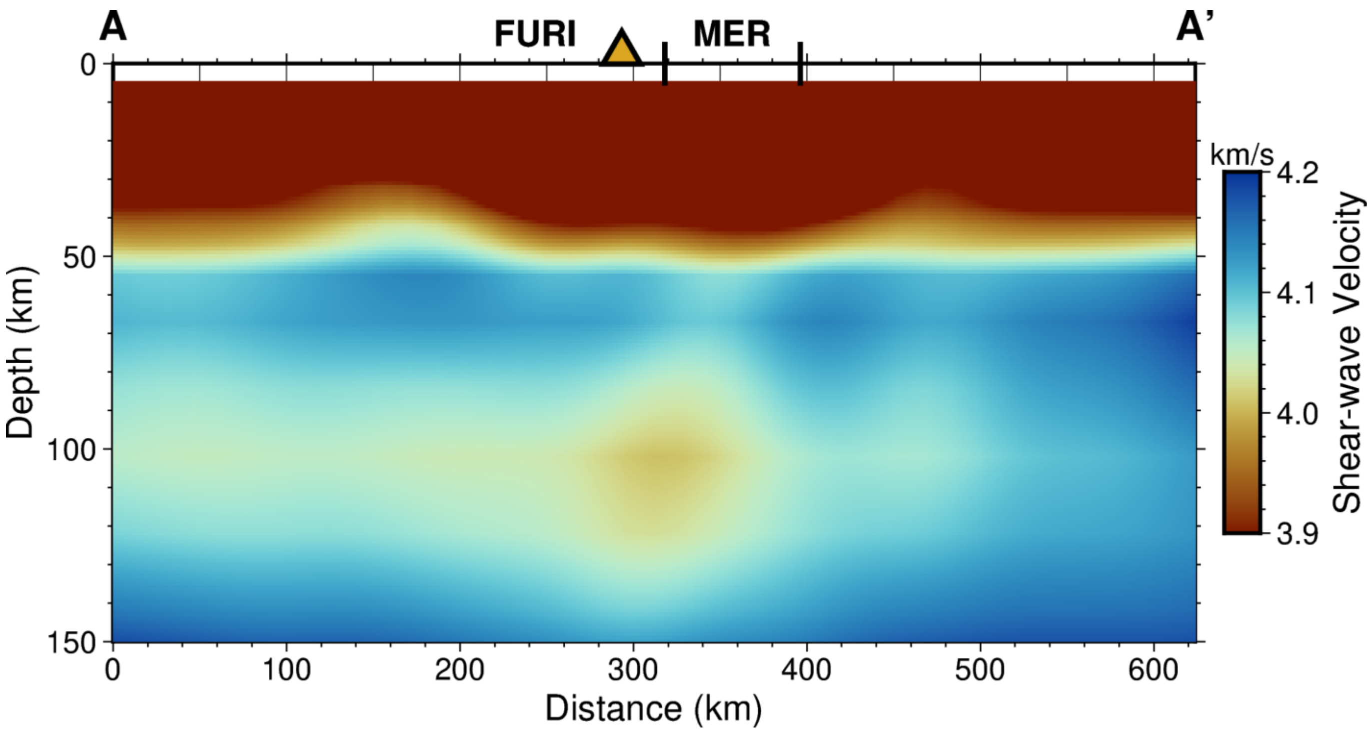


Figure 16 Interpolated cross-section through the shear-wave velocity model of Chambers et al. (2022). This cross-section is approximately perpendicular to the Main Ethiopia Rift and passes through FURI. The location and start/end points of the section (A-A') are shown in Figure 10. Black vertical lines indicate the approximate location of the Main Ethiopia Rift along the cross section. To reveal anomalies in the upper mantle, the colour scale is clipped at $3.9\,\mathrm{km\,s^{-1}}$ which masks crustal features. For details of crustal features which can be seen in the tomography, readers should refer to Chambers et al. (2022).

With further instrumentation and improvement of rock physics constraints, it may be possible to constrain the properties of fluid-filled fractures at a range of length scales within the Earth.

Acknowledgements

This work was supported by the Natural Environment Research Council [NE/S010203/1]. We would like to thank Robert Nowack and Rasheed Ajala for their insightful reviews. We would also like to thank the editor, Suzan van der Lee, for additional comments which improved the manuscript. We are grateful to Emma Chambers for sharing details of the tomography model used in this study. The facilities of EarthScope Consortium were used for access to waveforms, related metadata, and/or derived products used in this study. These services are funded through the Seismological Facility for the Advancement of Geoscience (SAGE) Award of the National Science Foundation under Cooperative Support Agreement EAR-1851048. Figures were produced using Obspy (Beyreuther et al., 2010), GMT version 6 as implemented in PyGMT version 0.9.0 (Wessel et al., 2019; Uieda et al., 2023) and matplotlib (Hunter, 2007).

Data and code availability

The data and code used in this article are available on Zenodo (Asplet et al., 2023).

Competing interests

The authors have no competing interests.

References

Abramson, E. H., Brown, J. M., Slutsky, L. J., and Zaug, J. The elastic constants of San Carlos olivine to 17 GPa. *Journal of Geophysical Research: Solid Earth*, 102(B6):12253–12263, 1997. doi: https://doi.org/10.1029/97JB00682.

Aki, K. and Richards, P. G. *Quantitative Seismology — Theory and Methods*, volume 1, chapter 5, pages 170–182. W.H. Freeman, 1980.

Albuquerque Seismological Laboratory/USGS. Global Seismograph Network (GSN - IRIS/USGS), 2014. doi: 10.7914/SN/IU.

Al-Harrasi, O. H., Kendall, J., and Chapman, M. Fracture characterization using frequency-dependent shear wave anisotropy analysis of microseismic data. *Geophysical Journal International*, 185(2):1059–1070, 2011. doi: 10.1111/j.1365-246x.2011.04997.x.

Asplet, J., Wookey, J., and Kendall, M. Inversion of shear wave waveforms reveal deformation in the lowermost mantle. *Geophysical Journal International*, 232(1):97–114, 2022. doi: 10.1093/gji/ggac328.

Asplet, J., Wookey, J., Kendall, M., Chapman, M., and Das, R. Supplementary material for Shear-wave attenuation anisotropy: a fluid detection tool, 2023. doi: 10.5281/zenodo.8275968.

Ayele, A., Stuart, G., and Kendall, J. Insights into rifting from shear wave splitting and receiver functions: an example from Ethiopia. *Geophysical Journal International*, 157(1):354–362, 2004. doi: 10.1111/j.1365-246x.2004.02206.x.

Backus, G. E. Long-wave elastic anisotropy produced by horizontal layering. *Journal of Geophysical Research*, 67(11):4427–4440, 1962. doi: 10.1029/jz067i011p04427.

Bacon, C. A., Johnson, J. H., White, R. S., and Rawlinson, N. On the Origin of Seismic Anisotropy in the Shallow Crust of the Northern Volcanic Zone, Iceland. *Journal of Geophysical Research: Solid Earth*, 127(1), 2022. doi: 10.1029/2021jb022655.

Baird, A. F., Kendall, J.-M., and Angus, D. A. Frequency-dependent seismic anisotropy due to fractures: Fluid flow versus scatteringFrequency-dependent anisotropy. *Geophysics*, 78 (2):WA111–WA122, 2013. doi: 10.1190/geo2012-0288.1.

Baird, A. F., Kendall, J.-M., Sparks, R. S. J., and Baptie, B. Transtensional deformation of Montserrat revealed by shear wave splitting. *Earth and Planetary Science Letters*, 425:179–186, 2015. doi: 10.1016/j.epsl.2015.06.006.

Barnes, A. E. Instantaneous spectral bandwidth and dominant frequency with applications to seismic reflection data. *Geophysics*, 58(3):419–428, 1993. doi: 10.1190/1.1443425.

Bastow, I. D., Nyblade, A. A., Stuart, G. W., Rooney, T. O., and Benoit, M. H. Upper mantle seismic structure beneath the Ethiopian hot spot: Rifting at the edge of the African low-velocity anomaly. *Geochemistry, Geophysics, Geosystems*, 9(12), 2008. doi: 10.1029/2008gc002107.

Bastow, I. D., Pilidou, S., Kendall, J.-M., and Stuart, G. W. Melt-induced seismic anisotropy and magma assisted rifting in Ethiopia: Evidence from surface waves. *Geochemistry, Geophysics, Geosystems*, 11(6), 2010. doi: https://doi.org/10.1029/2010GC003036.

Best, A. I., Sothcott, J., and McCann, C. A laboratory study of seismic velocity and attenuation anisotropy in near-surface sedimentary rocks. *Geophysical Prospecting*, 55(5):609–625, 2007. doi: https://doi.org/10.1111/j.1365-2478.2007.00642.x.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., and Wassermann, J. ObsPy: A Python Toolbox for Seismology. *Seismological Research Letters*, 81(3):530–533, 2010. doi: 10.1785/gssrl.81.3.530.

Buck, W. R. *1. Consequences of Asthenospheric Variability on Continental Rifting*, pages 1–30. Columbia University Press, New York Chichester, West Sussex, 2004. doi: doi:10.7312/karn12738-002.

Carter, A. J. and Kendall, J. Attenuation anisotropy and the relative frequency content of split shear waves. *Geophysical Journal International*, 165(3):865–874, 2006. doi: 10.1111/j.1365-246x.2006.02929.x.

Červenỳ, V., Molotkov, I. A., and Pšenčík, I. *Ray method in seismology*, chapter 2, pages 47–50. Universita Karlova, 1977.

Chambers, E. L., Harmon, N., Keir, D., and Rychert, C. A. Using Ambient Noise to Image the Northern East African Rift. *Geochemistry, Geophysics, Geosystems*, 20(4):2091–2109, 2019. doi: 10.1029/2018gc008129.

Chambers, E. L., Harmon, N., Rychert, C. A., Gallacher, R. J., and Keir, D. Imaging the seismic velocity structure of the crust and upper mantle in the northern East African Rift using Rayleigh wave tomography. *Geophysical Journal International*, 230(3): ggac156–, 2022. doi: 10.1093/gji/ggac156.

Chapman, M. Frequency-dependent anisotropy due to meso-scale fractures in the presence of equant porosity. *Geophysical Prospecting*, 51(5):369–379, 2003. doi: 10.1046/j.1365-2478.2003.00384.x.

Chapman, M., Maultzsch, S., and Liu, E. Some Estimates of the Squirt-flow Frequency. volume All Days of *SEG International Exposition and Annual Meeting*, pages SEG–2003–1290, 10 2003.

Chevrot, S. Multichannel analysis of shear wave splitting. *Journal of Geophysical Research: Solid Earth*, 105(B9):21579–21590, 2000. doi: 10.1029/2000jb900199.

Crampin, S. A review of wave motion in anisotropic and cracked elastic-media. *Wave Motion*, 3(4):343–391, 1981. doi: 10.1016/0165-2125(81)90026-3.

Crampin, S. Effective anisotropic elastic constants for wave propagation through cracked solids. *Geophysical Journal of the Royal Astronomical Society*, 76(1):135–145, 1984. doi: 10.1111/j.1365-246x.1984.tb05029.x.

Dasios, A., Astin, T., and McCann, C. Compressional-wave Q estimation from full-waveform sonic data. *Geophysical Prospecting*, 49(3):353–373, 2001. doi: 10.1046/j.1365-2478.2001.00259.x.

Durand, S., Matas, J., Ford, S., Ricard, Y., Romanowicz, B., and Montagner, J.-P. Insights from ScS–S measurements on deep mantle attenuation. *Earth and Planetary Science Letters*, 374:101–110, 2013. doi: 10.1016/j.epsl.2013.05.026.

Engelhard, L. Determination of Seismic-Wave Attenuation By Complex Trace Analysis. *Geophysical Journal International*, 125(2): 608–622, 1996. doi: 10.1111/j.1365-246x.1996.tb00023.x.

Eshelby, J. D. The determination of the elastic field of an ellipsoidal inclusion, and related problems. *Proceedings of the royal society of London. Series A. Mathematical and physical sciences*, 241 (1226):376–396, 1957.

Eshetu, A., Mammo, T., and Tilmann, F. Imaging the Ethiopian Rift Region Using Transdimensional Hierarchical Seismic Noise Tomography. *Pure and Applied Geophysics*, 178(11):4367–4388, 2021. doi: 10.1007/s00024-021-02880-2.

Ford, H. A., Goldhagen, G., Byrnes, J. S., and Brounce, M. N. New Insight into the Physical Properties of the East African Mantle from Seismic Attenuation. In *AGU Fall Meeting Abstracts*, volume 2022, pages T43B–03, Dec. 2022.

Ford, S. R., Garnero, E. J., and Thorne, M. S. Differential t* measurements via instantaneous frequency matching: observations of lower mantle shear attenuation heterogeneity beneath western Central America. *Geophysical Journal International*, 189(1): 513–523, 2012. doi: 10.1111/j.1365-246x.2011.05348.x.

Futterman, W. I. Dispersive body waves. *Journal of Geophysical Research (1896-1977)*, 67(13):5279–5291, 1962. doi: 10.1029/JZ067i013p05279.

Gabor, D. Theory of Communication. *Journal of the Institution of Electrical Engineers*, 93(26):429–441, July 1946. http://genesis.eecg.toronto.edu/gabor1946.pdf.

Galvin, R. J. and Gurevich, B. Effective properties of a poroelastic medium containing a distribution of aligned cracks. *Journal of Geophysical Research: Solid Earth*, 114(B7), 2009. doi: 10.1029/2008jb006032.

Galvin, R. J. and Gurevich, B. Frequency-dependent anisotropy of porous rocks with aligned fractures. *Geophysical Prospecting*, 63(1):141–150, 2015. doi: 10.1111/1365-2478.12177.

Hall, S. A., Kendall, J. M., and Baan, M. v. d. Some comments on the effects of lower-mantle anisotropy on SKS and SKKS phases. *Physics of the Earth and Planetary Interiors*, 146(3-4):469–481, 2004. doi: 10.1016/j.pepi.2004.05.002.

Hammond, J. O. S., Kendall, J., Wookey, J., Stuart, G. W., Keir, D., and Ayele, A. Differentiating flow, melt, or fossil seismic anisotropy beneath Ethiopia. *Geochemistry, Geophysics, Geosystems*, 15(5):1878–1894, 2014. doi: 10.1002/2013gc005185.

Hammond, W. C. and Humphreys, E. D. Upper mantle seismic wave attenuation: Effects of realistic partial melt distribution. *Journal of Geophysical Research: Solid Earth*, 105(B5):10987–10999, 2000. doi: https://doi.org/10.1029/2000JB900042.

Holtzman, B. K. and Kendall, J. Organized melt, seismic anisotropy, and plate boundary lubrication. *Geochemistry, Geophysics, Geosystems*, 11(12):n/a–n/a, 2010. doi:

10.1029/2010gc003296.

Hudson, J. A. Overall properties of a cracked solid. *Mathematical Proceedings of the Cambridge Philosophical Society*, 88(2): 371–384, 1980. doi: 10.1017/s0305004100057674.

Hudson, J. A. Wave speeds and attenuation of elastic waves in material containing cracks. *Geophysical Journal of the Royal Astronomical Society*, 64(1):133–150, 1981. doi: 10.1111/j.1365-246x.1981.tb02662.x.

Hudson, J. A., Liu, E., and Crampin, S. The mechanical properties of materials with interconnected cracks and pores. *Geophysical Journal International*, 124(1):105–112, 1996. doi: 10.1111/j.1365-246x.1996.tb06355.x.

Hudson, T. S., Asplet, J., and Walker, A. Automated shear-wave splitting analysis for single- and multi- layer anisotropic media. *Seismica*, 6 2023. doi: 10.31223/x5r67z.

Hunter, J. D. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(3):90–95, 2007. doi: 10.1109/MCSE.2007.55.

International Seismological Centre. ISC Bulletin. 2023. doi: https://doi.org/10.31905/D808B830.

Jakobsen, M., Johansen, T. A., and McCann, C. The acoustic signature of fluid flow in complex porous media. *Journal of Applied Geophysics*, 54(3-4):219–246, 2003. doi: 10.1016/j.jappgeo.2002.11.004.

Jin, Z., Chapman, M., and Papageorgiou, G. Frequency-dependent anisotropy in a partially saturated fractured rock. *Geophysical Journal International*, 215(3):1985–1998, 2018. doi: 10.1093/gji/ggy399.

Kendall, J.-M. Seismic anisotropy in the boundary layers of the mantle. In Karato, S., Forte, A., Liebermann, R., Masters, G., and Stixtrude, L., editors, *Earth's Deep Interior: Mineral physics and Tomography From the Atomic to the Global Scale*, volume 117 of *Geophysical Monographs*, pages 133–159. American Geophysical Union, 2000. doi: 10.1029/GM117p0133.

Kendall, J.-M., Stuart, G. W., Ebinger, C. J., Bastow, I. D., and Keir, D. Magma-assisted rifting in Ethiopia. *Nature*, 433(7022):146–148, 2005. doi: 10.1038/nature03161.

Lawrence, J. F. and Wysession, M. E. QLM9: A new radial quality factor (Qμ) model for the lower mantle. *Earth and Planetary Science Letters*, 241(3-4):962–971, 2006. doi: 10.1016/j.epsl.2005.10.030.

Li, Z., Leng, K., Jenkins, J., and Cottaar, S. Kilometer-scale structure on the core–mantle boundary near Hawaii. *Nature Communications*, 13(1):2787, 2022. doi: 10.1038/s41467-022-30502-5.

Liu, C. and Grand, S. P. Seismic attenuation in the African LLSVP estimated from PcS phases. *Earth and Planetary Science Letters*, 489:8–16, 2018. doi: 10.1016/j.epsl.2018.02.023.

Liu, E., Chapman, M., Varela, I., Li, X., Queen, J. H., and Lynn, H. Velocity and attenuation anisotropy Implication of seismic fracture characterizations. *The Leading Edge*, 26(9):1170–1174, 2007. doi: 10.1190/1.2780788.

Liu, J., Li, J., Hrubiak, R., and Smith, J. S. Origins of ultralow velocity zones through slab-derived metallic melt. *Proceedings of the National Academy of Sciences*, 113(20):5547–5551, 2016a. doi: 10.1073/pnas.1519540113.

Liu, Z., Park, J., and Karato, S. Seismological detection of low-velocity anomalies surrounding the mantle transition zone in Japan subduction zone. *Geophysical Research Letters*, 43(6): 2480–2487, 2016b. doi: 10.1002/2015gl067097.

Mainprice, D. Seismic anisotropy of the deep Earth from a mineral and rock physics perspective. Treatise of Geophysics, 2. In Schubert, G., editor, *Treatise of Geophysics*, volume 2, pages 437–491. Elsevier, 1st edition, 2015. doi: 10.1016/B978-044452748-6.00045-6.

Matheney, M. P. and Nowack, R. L. Seismic attenuation values obtained from instantaneous-frequency matching and spectral ratios. *Geophysical Journal International*, 123(1):1–15, 1995. doi: 10.1111/j.1365-246x.1995.tb06658.x.

Mavko, G. and Nur, A. Melt squirt in the asthenosphere. *Journal of Geophysical Research (1896-1977)*, 80(11):1444–1448, 1975. doi: https://doi.org/10.1029/JB080i011p01444.

Muller, G. Rheological properties and velocity dispersion of a medium with power-law dependence of Q on frequency. *Journal of Geophysics*, 54(1):20–29, 1984.

Quan, Y. and Harris, J. M. Seismic attenuation tomography using the frequency shift method. *Geophysics*, 62(3):895–905, 1997.

Restivo, A. and Helffrich, G. Teleseismic shear wave splitting measurements in noisyenvironments. *Geophysical Journal International*, 137(3):821–830, 1999. doi: 10.1046/j.1365-246x.1999.00845.x.

Rubino, J. G. and Holliger, K. Seismic attenuation and velocity dispersion in heterogeneous partially saturated porous rocks. *Geophysical Journal International*, 188(3):1088–1102, 2012. doi: 10.1111/j.1365-246x.2011.05291.x.

Rychert, C. A., Hammond, J. O. S., Harmon, N., Kendall, J. M., Keir, D., Ebinger, C., Bastow, I. D., Ayele, A., Belachew, M., and Stuart, G. Volcanism in the Afar Rift sustained by decompression melting with minimal plume influence. *Nature Geoscience*, 5(6): 406–409, 2012. doi: 10.1038/ngeo1455.

Saha, J. Relationship Between Fourier And Instantaneous Frequency. *1987 SEG Annual Meeting*, pages SEG–1987–0591, 1987.

Schlaphorst, D., Silveira, G., Mata, J., Krüger, F., Dahm, T., and Ferreira, A. M. G. Heterogeneous seismic anisotropy beneath Madeira and Canary archipelagos revealed by local and teleseismic shear wave splitting. *Geophysical Journal International*, 233 (1):510–528, 2022. doi: 10.1093/gji/ggac472.

Schmandt, B., Jacobsen, S. D., Becker, T. W., Liu, Z., and Dueker, K. G. Dehydration melting at the top of the lower mantle. *Science*, 344(6189):1265–1268, 2014. doi: 10.1126/science.1253358.

Shearer, P. M. *Introduction to Seismology*. Cambridge University Press, 2019. doi: 10.1017/9781316877111.

Silver, P. G. and Chan, W. W. Implications for continental strucutre and evolution from seismic anisotropy. *Nature*, 331(6185):450, 1988. doi: 10.1038/335034a0.

Silver, P. G. and Chan, W. W. Shear wave splitting and subcontinental mantle deformation. *Journal of Geophysical Research*, 96(B10):16429–16454, 1991. doi: 10.1029/91jb00899.

Solazzi, S. G., Lissa, S., Rubino, J. G., and Holliger, K. Squirt flow in partially saturated cracks: a simple analytical model. *Geophysical Journal International*, 227(1):680–692, 2021. doi: 10.1093/gji/ggab249.

Sun, Y., Carcione, J. M., and Gurevich, B. Squirt-flow seismic dispersion models: a comparison. *Geophysical Journal International*, 222(3):2068–2082, 2020. doi: 10.1093/gji/ggaa274.

Taner, M. T., Koehler, F., and Sheriff, R. E. Complex seismic trace analysis. *Geophysics*, 44(6):1041–1063, 1979. doi: 10.1190/1.1440994.

Teanby, N. A., Kendall, J., and Baan, M. V. D. Automation of Shear-Wave Splitting Measurements using Cluster Analysis. *Bulletin of the Seismologial Society of America*, 94(2):453–463, 2004. doi: 10.1785/0120030123.

Thomsen, L. Elastic anisotropy due to aligned cracks in porous rock. *Geophysical Prospecting*, 43(6):805–829, 1995. doi: 10.1111/j.1365-2478.1995.tb00282.x.

Tod, S. The effects on seismic waves of interconnected nearly aligned cracks. *Geophysical Journal International*, 146(1): 249–263, 2001. doi: 10.1046/j.1365-246x.2001.00451.x.

Uieda, L., Tian, D., Leong, W. J., Schlitzer, W., Grund, M., Jones, M., Fröhlich, Y., Toney, L., Yao, J., Magen, Y., Jing-Hui, T., Materna, K., Belem, A., Newton, T., Anant, A., Ziebarth, M., Quinn, J., and Wessel, P. PyGMT: A Python interface for the Generic Mapping Tools. doi: 10.5281/zenodo.7772533. The development of PyGMT has been supported by NSF grants OCE-1558403 and EAR-1948603.

Usher, P. J., Kendall, J., Kelly, C. M., and Rietbrock, A. Measuring changes in fracture properties from temporal variations in anisotropic attenuation of microseismic waveforms. *Geophysical Prospecting*, 65(S1):347–362, 2017. doi: 10.1111/1365-2478.12551.

Verdon, J. P. and Kendall, J.-M. Detection of multiple fracture sets using observations of shear-wave splitting in microseismic data. *Geophysical Prospecting*, 59(4):593–608, 2011. doi: 10.1111/j.1365-2478.2010.00943.x.

Walker, A. M. and Wookey, J. MSAT—A new toolkit for the analysis of elastic and seismic anisotropy. *Computers & Geosciences*, 49: 81–90, 2012. doi: 10.1016/j.cageo.2012.05.031.

Walsh, E., Arnold, R., and Savage, M. K. Silver and Chan revisited. *Journal of Geophysical Research: Solid Earth*, 118(10): 5500–5515, 2013. doi: 10.1002/jgrb.50386.

Wenzlau, F., Altmann, J. B., and Müller, T. M. Anisotropic dispersion and attenuation due to wave-induced fluid flow: Quasi-static finite element modeling in poroelastic solids. *Journal of Geophysical Research: Solid Earth*, 115(B7), 2010. doi: 10.1029/2009jb006644.

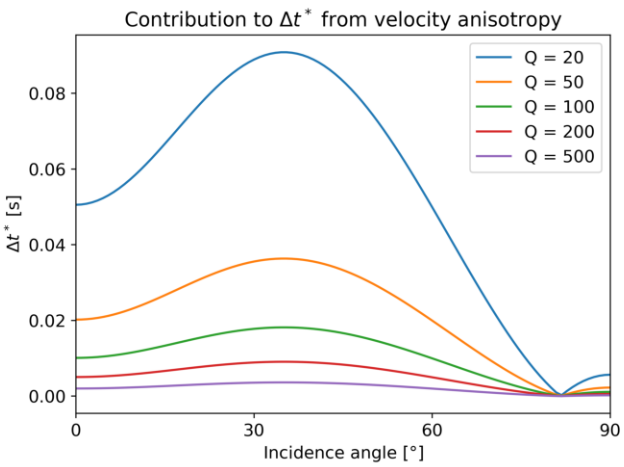
Werner, U. and Shapiro, S. A. Frequency-dependent shear-wave splitting in thinly layered media with intrinsic anisotropy. *Geophysics*, 64(2):604–608, 1999. doi: https://doi.org/10.1190/1.1444567.

Wessel, P., Luis, J. F., Uieda, L., Scharroo, R., Wobbe, F., Smith, W. H. F., and Tian, D. The Generic Mapping Tools Version 6. *Geochemistry, Geophysics, Geosystems*, 20(11):5556–5564, 2019. doi: https://doi.org/10.1029/2019GC008515.

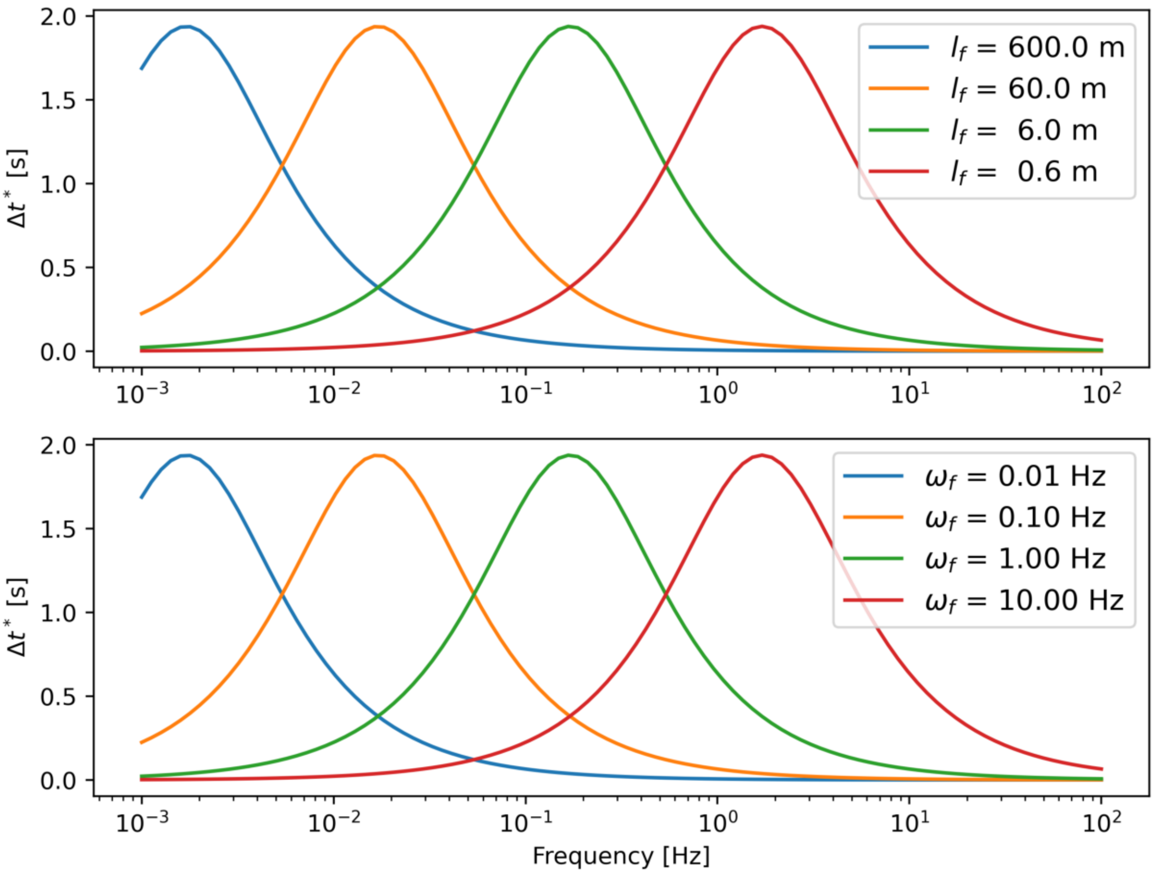
Whaler, K. and Hautot, S. The electrical resistivity structure of the crust beneath the northern Main Ethiopian Rift. *Geological Society, London, Special Publications*, 259(1):293–305, 2006. doi: 10.1144/GSL.SP.2006.259.01.22.

Wuestefeld, A., Al-Harrasi, O., Verdon, J. P., Wookey, J., and Kendall, J. M. A strategy for automated analysis of passive microseismic data to image seismic anisotropy and fracture characteristics. *Geophysical Prospecting*, 58(5):755–773, 2010. doi: 10.1111/j.1365-2478.2010.00891.x.

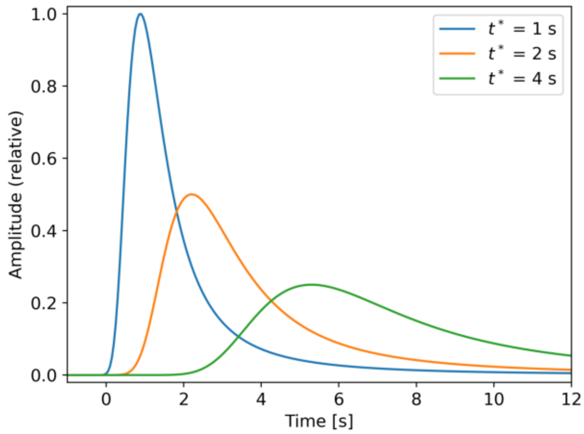
Zhubayev, A., Houben, M. E., Smeulders, D. M. J., and Barnhoorn, A. Ultrasonic velocity and attenuation anisotropy of shales, Whitby, United Kingdom. *Geophysics*, 81(1):D45–D56, 2016. doi: 10.1190/geo2015-0211.1.



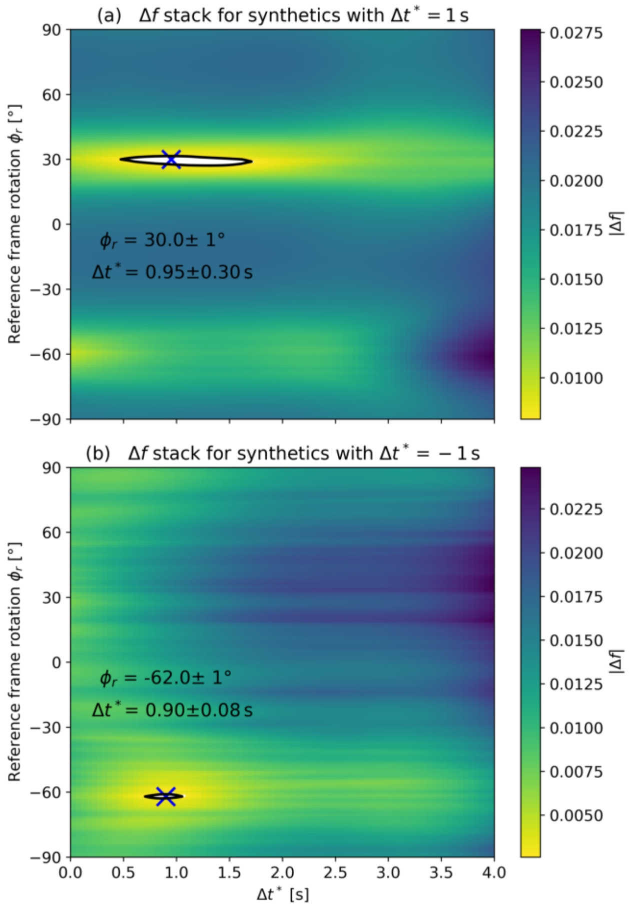
Supplementary Figure 1 Attenuation anisotropy, Δt^*, calculated using an elastic tensor for single-crystal olivine (Abramson et al., 1997) taken from the MSAT toolkit (Walker and Wookey, 2012) and assuming a $50\,\mathrm{km}$ path length. Δt^* is calculated for a range assumed isotropic Q values, where the only contribution to Δt^* in equation 3 is the velocity anisotropy obtained from solving the Christoffel equation for the elastic tensor.



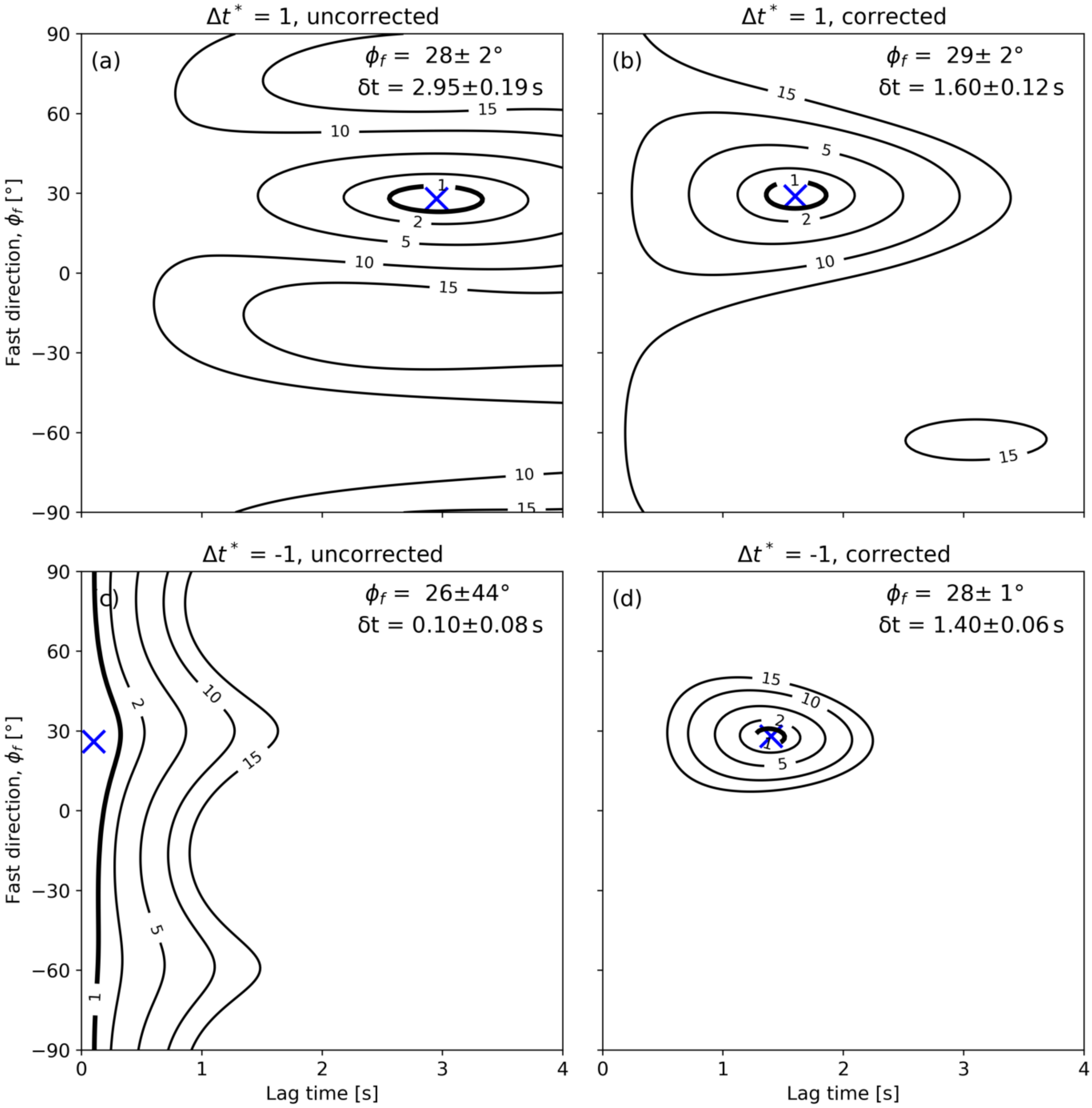
Supplementary Figure 2 Δt^* calculated using the squirt flow model as a function of frequency. We calculate Δt^* for a range of fracture lengths l_f (top panel) and convert these to representative fracture-scale squirt flow frequencies using equation 10. Here we can see that different length scale fractures will induce a squirt-flow response (and attenuation anisotropy) in different frequency bands.



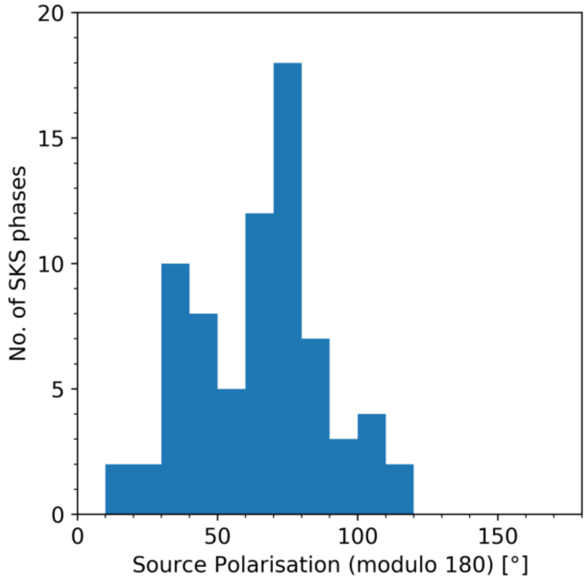
Supplementary Figure 3 Pulse shapes obtained when attenuating a delta function at $t = 0$ for a reference frequency of $10\,\mathrm{Hz}$. Adapted after Shearer (2019).



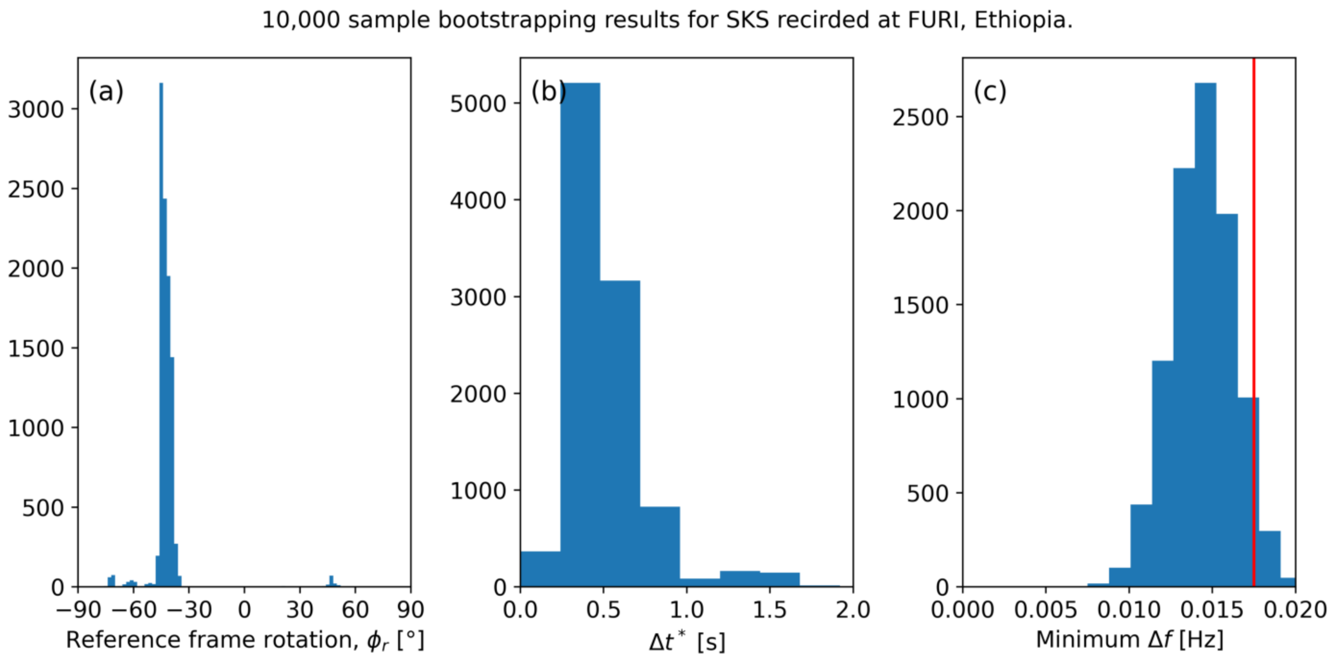
Supplementary Figure 4 Source polarisation weighted, stacked $|\Delta f|$ surfaces. Each panel shows the $|\Delta f|$ stack measured for 100 Gabor wavelet synthetics generated with shear-wave splitting parameters $\phi_f \sim \mathcal{N}(30, 5), \delta t \sim \mathcal{N}(1.5, 0.15)$ and $\Delta t^* = 1\,\mathrm{s}$ (a) or $\Delta t^* = -1\,\mathrm{s}$ (b). Each synthetic is generated with a random source polarisation and with a dominant frequency drawn from $f \sim \mathcal{N}(0.1, 0.02)$. Surfaces are drawn following Figure 7.



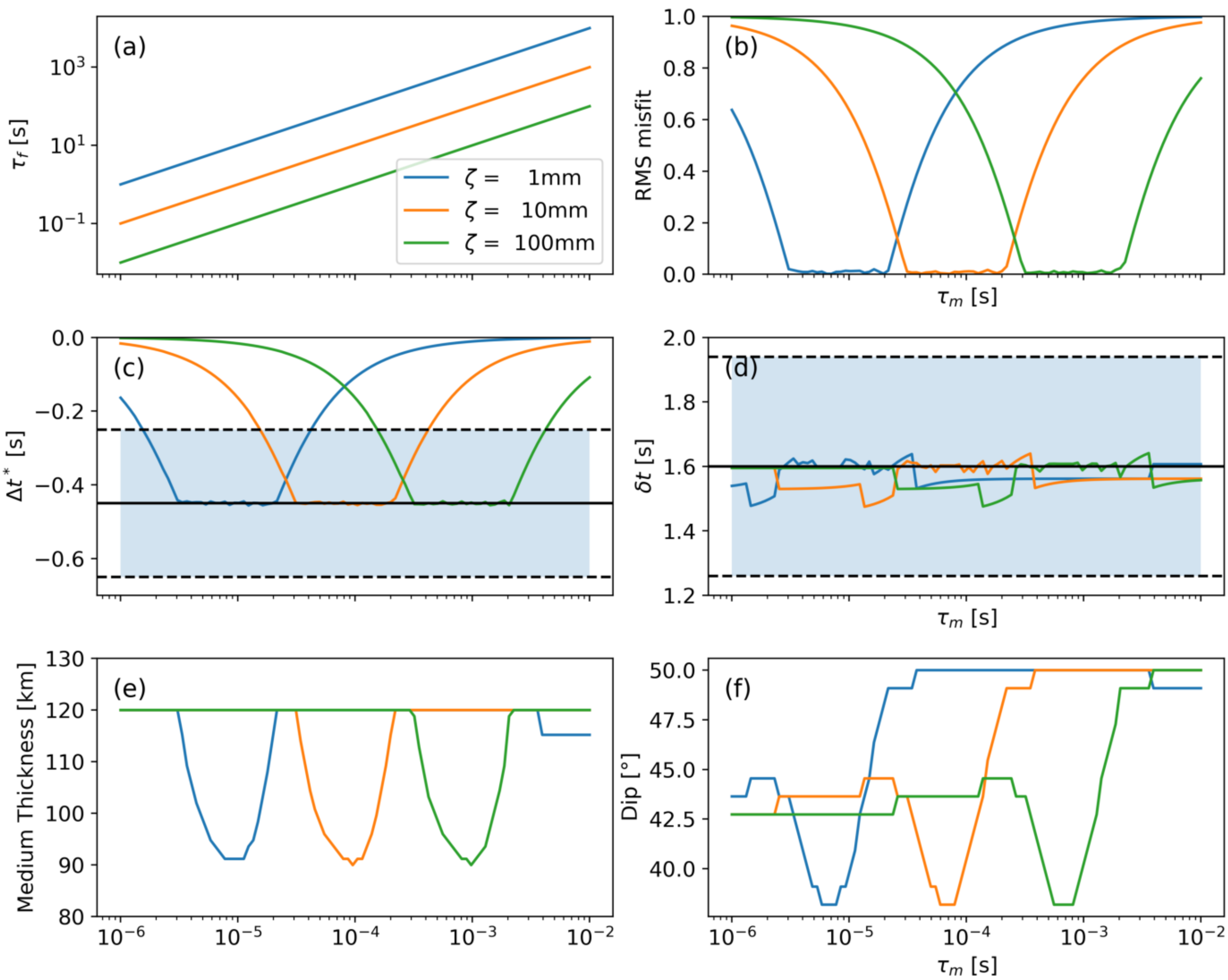
Supplementary Figure 5 Results of synthetic shear-wave splitting measurement stacking, following the method of Restivo and Helffrich (1999). We generate 100 synthetics where $\phi_f \sim \mathcal{N}(30, 5)$ and $\delta t \sim \mathcal{N}(1.5, 0.15)$. Attenuation anisotropy of $\Delta t^* = 1\,\mathrm{s}$ (a,b) or $\Delta t^* = -1\,\mathrm{s}$ (c,d) is applied. Panels (a,c) show the shear-wave splitting results if we do not correct for this attenuation anisotropy. Panels (c,d) show the result after we correct the synthetic data using measurements of $\phi_r, \Delta t^*$ made using our source polarisation stacking method. The stacked λ_2 surfaces are normalised by the 95% confidence value, indicated by the bold contours, which is derived from an F-test (Silver and Chan, 1991; Restivo and Helffrich, 1999).



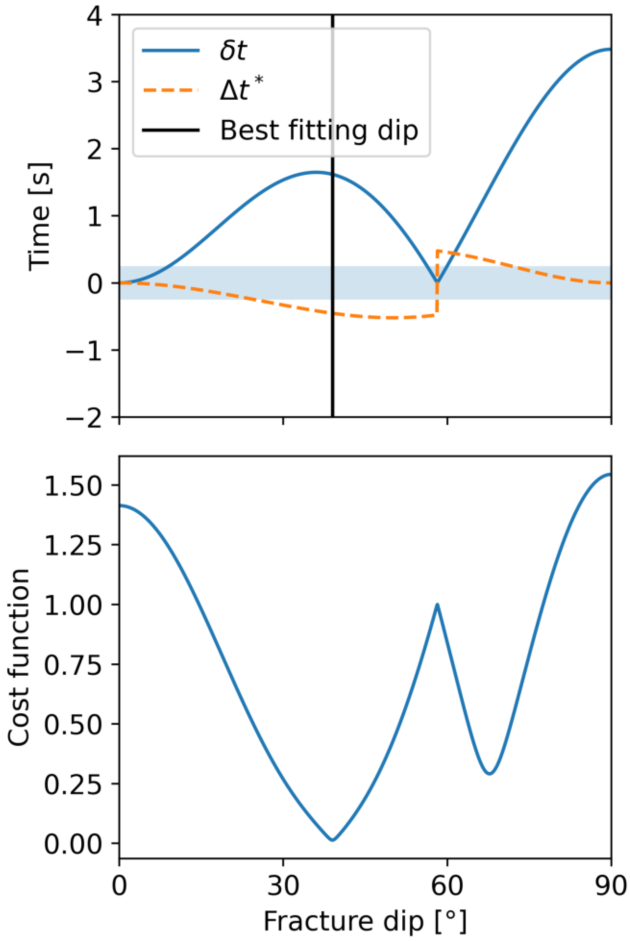
Supplementary Figure 6 Histogram showing the measured source polarisation (modulo $180°$) of the 74 SKS phases used in the shear-wave splitting and attenuation anisotropy measurements. Source polarisations are binned in intervals of $10°$, the same bins used in the source polarisation weighting when stacking the individual shear-wave splitting and attenuation anisotropy measurements. The achieved source polarisation coverage here is reasonable, ranging from $10°$ to $120°$, but is far from an ideal uniform distribution.



Supplementary Figure 7 Bootstrapped summary statistics for the $|\Delta f|$ measurement stacking for SKS data recorded at FURI, Ethiopia. Histograms show the parameters ϕ_r, (a) and Δt^* (b) along with the minimum $|\Delta f|$ of each bootstrapped stack (c). We draw 10,000 bootstrap samples, with replacement, from the 74 SKS phases used.



Supplementary Figure 8 Results of modelling the fracture dip and medium thickness, or ray path length, which best explain the observed δt and Δt^* at FURI, Ethiopia using a squirt flow model (Chapman, 2003). Due to the lack of constrain in the mineral-scale relaxation time, τ_m, we search over a range of τ_m values for assumed grain sizes of $1\,\mathrm{mm}$, $10\,\mathrm{mm}$ and $100\,\mathrm{mm}$ and fracture length of $1000\,\mathrm{m}$ (green).



Supplementary Figure 9 Modelling of δt and Δt^* as a function of fracture dip with a squirt flow model using the parameters in Table 1, $\tau_m = 9.55 \times 10^{-5}\,\mathrm{s}$ and a medium thickness of $90\,\mathrm{km}$. The top panel shows the modelled δt and Δt^*, in which the black line indicated the fracture dip which best fits the measured values. The shaded blue region shows the uncertainty in Δt^* measured at FURI $(\pm 0.2\,\mathrm{s})$ to indicate the space which could be plausibly fit by no attenuation anisotropy. The bottom panel shows the normalised least-square cost function used to find the best-fitting dip angle.

doi:10.26443/seismica.v3i1.1092

SEISMICA

Insights on the dip of fault zones in Southern California from modeling of seismicity with anisotropic point processes

Zachary E. Ross * [1]

[1]Seismological Laboratory, California Institute of Technology, Pasadena, CA, USA 91125

Abstract Accurate models of fault zone geometry are important for scientific and hazard applications. While seismicity can provide high-resolution point measurements of fault geometry, extrapolating these measurements to volumes may involve making strong assumptions. This is particularly problematic in distributed fault zones, which are commonly observed in immature faulting regions. In this study, we focus on characterizing the dip of fault zones in Southern California with the goal of improving fault models. We introduce a novel technique from spatial point process theory to quantify the orientation of persistent surficial features in seismicity, even when embedded in wide shear zones. The technique makes relatively mild assumptions about fault geometry and is formulated with the goal of determining the dip of a fault zone at depth. The method is applied to 11 prominent seismicity regions in Southern California. Overall, the results compare favorably with the geometry models provided by the SCEC Community Fault Model and other focused regional studies. More specifically, we find evidence that the Southern San Andreas and San Jacinto fault zones are both northeast dipping at seismogenic depths at the length scales of 1.0–4.0 km. In addition, we find more limited evidence for some depth dependent variations in dip that suggest a listric geometry. The developed technique can provide an independent source of information from seismicity to augment existing fault geometry models.

Production Editor:
Gareth Funning
Handling Editor:
Alice-Agnes Gabriel
Copy & Layout Editor:
Abhineet Gupta

Received:
August 03, 2023
Accepted:
May 13, 2024
Published:
June 18, 2024

1 Introduction

The geometrical properties of fault zones are basic, yet fundamental quantities in earthquake science. Earthquake rupture simulations need fault geometry models that faithfully capture these attributes in order to adequately quantify expected seismic hazard with physics-based approaches (Shaw et al., 2018; Rodgers et al., 2019; Melgar et al., 2016). Fault zones are the locus of intense deformation processes spanning a wide range of strain rates and contain valuable information on the long term history of these processes (Ben-Zion and Sammis, 2003); the geometry of a fault zone at a range of length scales, including any depth-dependent variations, can aid in reconstructing this history and constraining the physical processes involved (Norris and Toy, 2014; Schulte-Pelkum et al., 2020).

A fault zone's geometry is commonly assessed from a variety of sources. These include focal mechanisms determined with seismological methods (Lin et al., 2007; Shelly et al., 2016), high-resolution seismicity catalogs (Chiaraluce et al., 2017; Ross et al., 2017a), various types of seismic imaging (Sato et al., 2005; Fuis et al., 2017; Lay et al., 2021; Bangs et al., 2023), geological data and mapped fault traces (Fletcher et al., 2014), and geodetic data (Lindsey and Fialko, 2013). These diverse information sources have their own uncertainties and sensitivities, making them complimentary when multiple sources are available; however it is not always straightforward to assimilate them. Several databases of fault geometry models have been produced with the goal of incorporating community consensus and providing established models with a documented provenance. These include faults at global scale (Bird, 2003; Hayes et al., 2012, 2018) and also some regional scales (Plesch et al., 2007, 2020b).

In this study, we aim to characterize the dip of fault zones in Southern California with high-resolution seismicity. We introduce a simple technique from the statistical field of spatial point processes that can measure fault zone dip independently from traditional methods, with the goal of augmenting the information available for constructing fault models. We first apply the method to four synthetic catalogs to demonstrate its suitability. We then apply the technique to eleven prominent seismicity regions across southern California to quantify the dip for different fault zone sections. These findings are compared with those of the SCEC Community Fault Model and other previous works in the area. We demonstrate that the method can reliably recover fault dip, including depth-dependent variations under some circumstances. Our primary scientific findings are that the San Jacinto and San Andreas fault zones appear to have significant northeasterly dips, whereas the Elsinore fault zone and Brawley Seismic Zone appear to be nearly vertical fault zones.

*Corresponding author: zross@caltech.edu

2 Methods

2.1 Preliminaries

Let $X \subset \mathbb{R}^D$ be a stochastic collection of points, i.e., a spatial point process (Daley and Vere-Jones, 2003). For a spatial domain $W \subset \mathbb{R}^D$, let $N(W)$ denote the number of points of X that are contained within W. For those readers familiar with measure theory, $N(\cdot)$ is a counting measure on W. Since X is a stochastic process, the mean number of points in W is given by the so-called intensity measure,

$$\Lambda(W) = \mathbb{E}\left(N(W)\right), \tag{1}$$

where $\mathbb{E}\left(\cdot\right)$ denotes an expected value. Let us also denote the volume of W in $\mathbb{R}^D$ as $|W|$. Then, for a stationary point process, the quantity $\lambda = \Lambda(W)/|W|$ is independent of the choice of W. While $\Lambda(W)$ describes the expected number of points within a particular fixed volume, it does not describe spatial correlation of event density, i.e., knowing $\Lambda(W)$ does not tell you anything about $\Lambda(V)$ for some other disjoint $V \subset \mathbb{R}^D$.

Instead, we need a different type of quantity to characterize the spatial correlation of points. For a typical point $u \in X$, one such choice is the K-function (Ripley, 1976),

$$K(r) = \frac{1}{\lambda}\mathbb{E}\left(\text{number of neighbors within radius } r \,|\, X \text{ has a point at } u\right). \tag{2}$$

The quantity $\lambda\, K(r)$ therefore quantifies the mean number of neighbors that any typical point will have within a sphere of radius r. The K-function is a cumulative function of r and was first introduced to seismology by Kagan and Knopoff (1980), where it is often referred to as a correlation integral; most commonly the K-function has been used to infer the fractal distribution of a set of hypocenters by fitting a power law to an empirical estimator of the K-function. A useful property of $K(r)$ is that it describes how point patterns are arranged in space, independently of the choice of W. This is because $K(r)$ is a second-order quantity and is analogous to a covariance, whereas $\Lambda(W)$ is a first-order quantity and is analogous to an expected value.

The function $K(r)$ has an inherent normalization property, which is seen by considering that for a Poisson process in 2D,

$$K_{pois}(r) = \pi r^2, \tag{3}$$

i.e. K_{pois} depends only on r (and not on λ). This is important as it allows for $K_{pois}(r)$ to be used as a reference, and if $K(r) > K_{pois}(r)$, it is said that X is clustered, since more of the points then locate within the sphere of radius r than expected for the equivalent Poisson process. This is only possible because K is conditional on a typical point existing at the center of the sphere.

The K-function can be estimated using the following empirical formula,

$$\hat{K}(r) = \frac{|W|}{m(m-1)}\sum_{i}^{m}\sum_{j\neq i}^{m}\mathbf{1}\{d_{ij} \leq r\}e_{ij}. \tag{4}$$

In this equation, $\mathbf{1}(\cdot)$ is the indicator function, d_{ij} is the Euclidean distance between points i and j, e_{ij} is an edge correction factor, m is the number of points in the observation window, and $|W|$ is the area (volume) of the observation window.

2.2 The cylindrical K-function

The K-function, as given above, is derived by assuming the point process is both stationary and isotropic, i.e. the likelihood of a point at u given a point at v depends only on the distance between them $r = \|u-v\|$. Seismicity, however exhibits strong spatial anisotropy at scales from local to global (Ross et al., 2022; Nasirzadeh et al., 2021; Møller and Toftaker, 2014; Rubin et al., 1999). Seismicity lineations, i.e., collections of hypocenters that align in the form of linear features, are commonly observed in the highest resolution catalogs (Gillard et al., 1996; Shearer, 2002). Sometimes, hypocenters align in the form of planar or surficial features (Ross et al., 2020; Cox, 2016). Both linear and planar seismicity features are evidence of anisotropic point patterns since the likelihood of a point at a location u given a point exists at v depends on not just the spatial separation between them, but also the orientation of the vector connecting them, i.e. $K = K(u - v)$.

Within the spatial statistics literature, there has been interest in detection and characterization of anisotropy in point processes (Møller and Toftaker, 2014; Møller et al., 2016; Safavimanesh and Redenbach, 2016; Nasirzadeh et al., 2021). One important development has been the cylindrical K-function (Møller et al., 2016), in which a cylinder is used in place of a sphere to characterize anisotropy that is effectively columnar. A cartoon example of this approach is shown in Figure 1, in which a cross section of seismicity is depicted. Here, the seismicity exhibits a dipping fabric that is orthogonal to the vector $\hat{n}$. When a cylinder defined by this normal vector is used (e.g., blue cylinder), the value of K is maximized, as the cylinder on average will enclose more points than a cylinder aligned with any other orientation (e.g., red cylinder). By computing K over all azimuths and polar angles, it is possible to detect anisotropy and quantify its orientation.

For a unit vector $n = [\cos\varphi\sin\theta, \sin\varphi\sin\theta, \cos\theta]$, let $C_n(r,t)$ denote a cylinder with radius r, height $2t$, and normal vector n. For an observed set of points, $\{x_1, ..., x_m\}$, the cylindrical K-function (Møller et al., 2016) is then computed as,

$$K_{cyl}(r,t,\theta,\varphi) = \frac{1}{\lambda^2}\sum_{i}^{m}\sum_{j\neq i}^{m}\mathbf{1}\{x_j - x_i \in C_n\}e_{ij}, \tag{5}$$

where the condition $x_j - x_i \in C_n$ is true if the vector separating x_j and x_i locates inside C_n, and e_{ij} is an edge correction factor. In this study, we use the translation-based edge correction, a routine choice in point processes in which the window W is translated by the vector $x_j - x_i$ and the amount of overlap between the translated window and the original window is computed,

$$e_{ij} = \frac{|W|}{|W \cap (W + x_j - x_i)|} \tag{6}$$

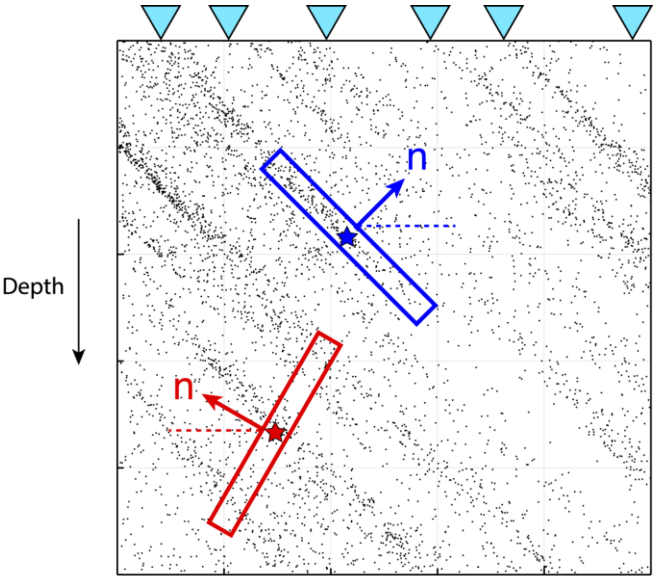


Figure 1 Illustration of method. A cylindrical K-function is computed by placing a disc with normal vector $\hat{n}$ centered on each event (stars). On average, the number of events contained in the disc is highest when the disc is aligned with seismicity lineations (blue box), resulting in a large value of K. Similarly, K is low when poorly aligned with seismicity lineations (red box). The best dip estimate is equal to the dip of $\hat{n}$ for which K is maximized. The method can detect dipping fabric even in distributed seismicity, such as in the cartoon, if a persistent orientation is present.

We propose K_{cyl} as a method to infer the dip of fault zones from seismicity, even when weakly localized as in Figure 1, due to these aforementioned properties. While Møller et al. (2016) focused on detecting columnar structures with K_{cyl} by using highly elongated cylinders (i.e., $r < t$), it can also be used to detect coherent surface-like structures in seismicity if the diameter of the cylinder is longer than its height (i.e., it is more aptly described as a disc, as in Figure 1). This disc-based formulation is the one we use in this study.

2.3 Demonstration with synthetic catalogs

We begin with four synthetically generated seismicity catalogs to demonstrate the method and provide additional insights into its usage. Furthermore, we use this opportunity to walk through the novel summary diagram used to visualize the results in this study.

Case 1: A single vertical planar fault. We randomly generate 1000 hypocenters drawn from a uniform distribution on a planar N–S trending vertical fault with a length of 50 km and seismogenic thickness of 20 km. We set $r = 0.1$ km, $t = 1.0$ km, and compute K_{cyl} on a grid with 2° spacing using equation 5. Figure 2 shows the seismicity in both map view and cross-section. It also shows K_{cyl} for this catalog in an upper-hemisphere stereographic projection, where the polar angle θ of the fault normal vector is given on the radial axis and the angle φ is given as the traditional azimuthal angle for such a diagram. Here, K_{cyl} correctly attains maxima at both $\varphi = 90°$ and $\varphi = 180°$, reflecting the symmetry of this particular dataset. The correct dip is also attained with little ambiguity.

Case 2: A single dipping planar fault. We randomly generate 1000 hypocenters drawn from a uniform distribution on a N–S striking 30° dipping planar fault with a length of 50 km and seismogenic thickness of 20 km. As with the previous example, K_{cyl} correctly recovers both the fault normal azimuth and the dip of the fault. Note that only one mode is present now in the K_{cyl} plot, as the break in symmetry leads to the other mode occurring in the lower hemisphere, and thus not in the plot.

Case 3: Distributed fault zone with vertical dip. We simulate seismicity occurring within a distributed fault zone having a vertical dip. Following the work of Møller et al. (2016), we choose 20 random vertical faults (with dimensions 50 km×20 km) that strike north-south. For each fault, we generate 500 random hypocenters that are then displaced randomly in the fault normal direction with Gaussian noise of 100 m to add complexity. The realization of this Poisson plane cluster process that we use is shown in Figure 2. K_{cyl} correctly identifies the same overall pattern as seen for the single planar vertical fault case, as there is just a single dominant orientation for the anisotropy even though 20 faults are present in the catalog. This demonstrates the potential for measuring fault dip even when the seismicity and fault zone is highly distributed, provided that the anisotropy is persistent across much of the seismicity.

Case 4: Distributed fault zone with conjugate faults. We simulate seismicity occurring within a distributed fault zone having conjugate faults with dips of around 45°. The dip is randomly perturbed so that not all angles are identically 45°. We create 20 faults that strike north-south, with half dipping to the west and half dipping to the east. For each fault, we randomly locate 500 hypocenters within it. The hypocenters are then displaced randomly in the fault normal direction with Gaussian noise of 100 m to add complexity. The resulting catalog is shown in Figure 2. K_{cyl} correctly indicates two orthogonally dipping faults with the same strike. This demonstrates the potential for measuring multiple fault dip angles, when present.

2.4 Application to Southern California seismicity

We now shift our focus to using K_{cyl} to quantify the dip for fault zones in Southern California. We use a high-resolution relocated seismicity catalog that covers the entirety of southern California and the northern part of Baja California for the period 1981–2019. The catalog used is based on the methodology of Hauksson et al. (2012) and has been updated for recent years (Figure 3). It contains 679,495 earthquakes that have been relocated with waveform cross-correlation, which form the highest quality subset. We focus only on the relocated events in this study. The catalog is publicly available from the Southern California Earthquake Data Cen-

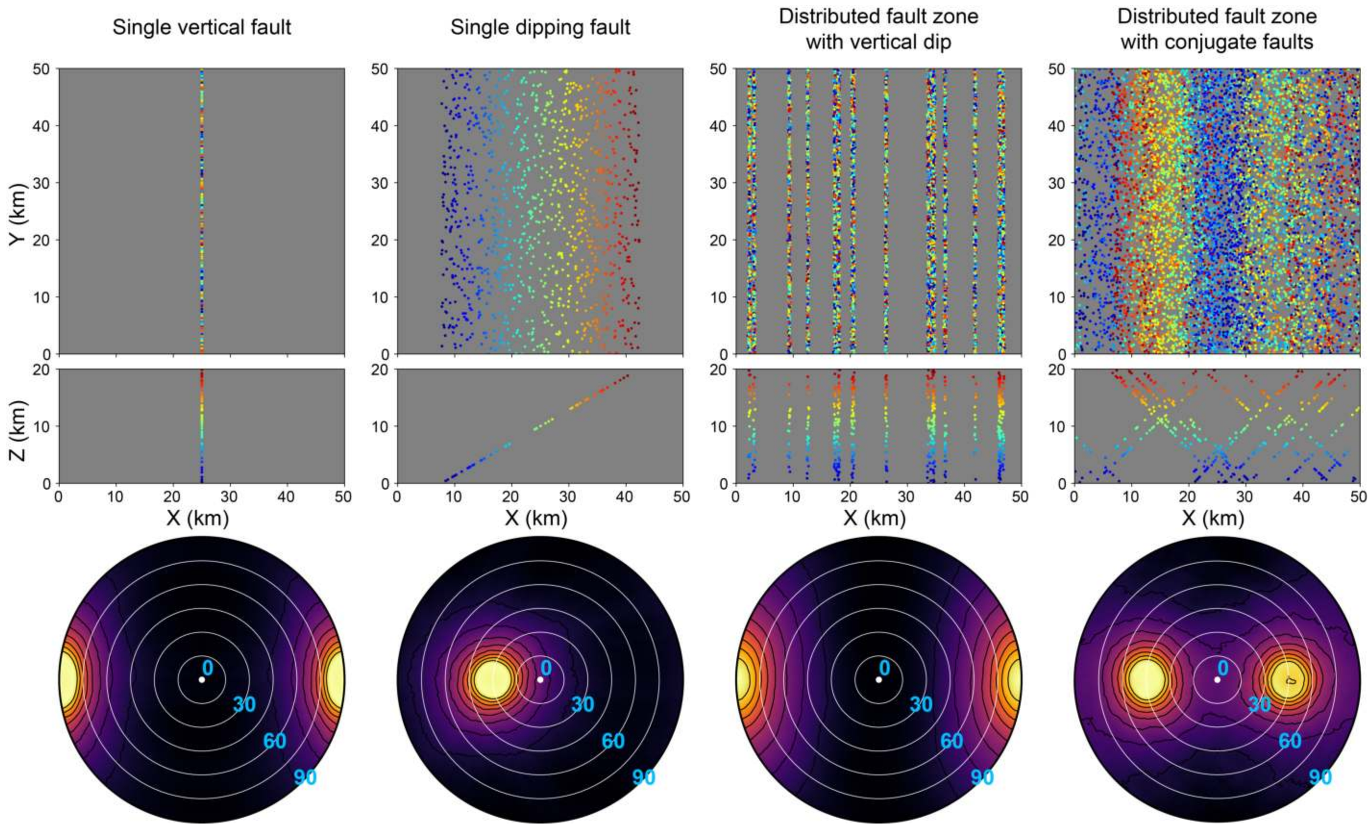


Figure 2 Method Demonstration with synthetic catalogs. Each column is a different seismicity catalog (described in main text). Events are colored by depth to enhance visibility. Upper row: map view of seismicity. Middle row: East-west cross section with seismicity projected onto it. For plotting purposes, seismicity is shown thinned by 95%. Bottom row: Stereographic projection of K_{cyl} for each catalog. Warmer colors indicate more intense clustering along a given fault normal azimuth and dip.

ter (Southern California Seismic Network, 2013). We use only the hypocenters and magnitudes for these catalogs.

We also considered using the the Quake Template Matching (QTM) catalog for southern California (Ross et al., 2019), which contains 10 times more events but spans only the period 2008–2017. Ultimately, we opted for the Hauksson et al. (2012) catalog because it is much longer in duration and the hypocenters are generally more precise; the many extra smaller events detected in the QTM catalog have fewer phase picks available and lead to an overall slight degradation in location accuracy as compared with the Hauksson et al. (2012) events, which is less desirable for this study.

For our analyses, we subset the catalog into 11 non-overlapping fault zone sections. They are denoted by red boxes in Figure 3 and described in more detail in Table 1; the number of earthquakes within each region is also given. These regions were chosen based on a variety of factors, including scientific or hazard importance, longstanding fault segment demarcation by the community, an abundance of seismicity, or clear geometrical boundaries. The list contains four sections of the San Jacinto Fault Zone, two sections of the San Andreas Fault Zone, four sections of the Elsinore Fault Zone, and the Brawley Seismic Zone. For all but one of the regions, there are thousands of earthquakes available, which is important to ensure the statistical estimators are robust.

Region #	Region Name	Number of Events
1	San Jacinto Fault Zone (Claremont)	14,340
2	San Jacinto Fault Zone (Hot Springs)	24,066
3	San Jacinto Fault Zone (Trifurcation Area)	29,914
4	San Jacinto Fault Zone (Borrego Mountain)	24,662
5	Southern San Andreas	723
6	San Gorgonio Pass	23,614
7	Brawley Seismic Zone	9,402
8	Elsinore Fault Zone (Whittier)	3,396
9	Elsinore Fault Zone (Julian)	17,644
10	Elsinore Fault Zone (Coyote Mountain)	6,864
11	Elsinore Fault Zone (Yuha)	21,939

Table 1 Description of the focus areas in Southern California

For each region, we compute K_{cyl} using the horizontal coordinates as defined in Figure 3 and using the depth range $[0, 22]$ km. We then use equation 5 to compute for three sets of parameters, (t, r) = (50 m, 500 m), (100 m, 1000 m), and (200 m, 2000 m). We compute K_{cyl} for $\theta \in [0, \pi]$ and $\varphi \in [0, 2\pi]$, i.e., the whole range possible, as we aim to estimate the dip of each fault zone without any prior knowledge. This framework also provides a means to perform hypothesis testing if several candidate scenarios for the dip are believed to be possible (which is covered in more detail in the discussion).

The domains for θ and φ are discretized with spacing of $1°$; this choice is mainly a balance between having sufficiently fine spatial resolution and computational efficiency, since the results are largely insensitive to them. Given K_{cyl}, the best estimate of the fault normal vector is defined by the values of θ and φ for which K_{cyl} is maximized (as in Figure 1). The best fault zone dip estimate is then $\delta = \pi - \theta$.

2.5 Dip uncertainty estimates

The polar diagrams for K_{cyl} are useful for visual examination of the results and identifying the most likely dip angle(s), but do not communicate the uncertainty associated with these measurements. To obtain uncertainty estimates, we use a bootstrapping approach designed for spatial resampling of these empirical estimators (Loh, 2008). We use this method to resample local K_{cyl} functions with replacement, compute an average K_{cyl} function for each bootstrap sample, measure $\delta = \pi - \theta$ corresponding to the peak of K_{cyl}, and repeat this process 1000 times. The ensemble of δ values resulting from the bootstrap procedure provides an estimate of the uncertainty.

2.6 Parameter selection and resolution

The two parameters t and r control the resolution of the method and here we give some additional insight and guidance around their usage. Generally speaking, it will be unknown beforehand what length scales are useful for measuring the dip. Thus, it is desirable to to compute K_{cyl} for a range of values. Figure 4 shows two schematic scenarios and the potential for resolving faults with the method. In Figure 4, a red disc of radius r and a blue disc of radius $2r$ are shown, with $t \ll r$ for both. In (a), the seismicity pattern has structure with an effective length scale of about $2r$. For this case, both the red and blue discs can resolve this anisotropy since the length scale is less than or equal to the diameter of the disc. Thus, the diameter of the disc is effectively an upper bound to the length scale of the anisotropy. In (b), the seismicity pattern exhibits a length scale comparable to the whole window. In this case, both the red and blue discs can resolve the anisotropy, however since both discs have a diameter smaller than the length scale of the seismicity, they are unable to provide information about larger length scales.

If the true hypocenter configuration exhibits planar anisotropy, then making the disc thickness t as small as possible will increase sensitivity for detecting anisotropy. However, the lower limit for whether t will be useful is closely related to the location errors in the respective direction. Thus, we recommend initially setting the value of t to be comparable to the estimated relative location error of most events.

Practically speaking, there will be limits to the value of r that can be used. The largest values of r used should depend on the dimensions of the spatial window, W; in particular, K_{cyl} will become unreliable as $2r$ approaches values of roughly 1/4 the shortest spatial dimension of W. This is true despite the use of an edge correction factor, as there will be little usable signal left to correct at these scales, similar to amplifying noise in seismic deconvolution. At the same time, r should still be much larger than t, in order to have sufficient sensitivity in detecting anisotropy. As the aspect ratio r/t approaches 1:1, K_{cyl} becomes effectively unable to identify anisotropy. Additionally, r should be large enough that enough events locate within the discs to constrain K_{cyl} to a desirable level (preferably as measured from the aforementioned bootstrap procedure).

For this study, we use a single fixed aspect ratio of $r/t = 10$, in part to simplify the process of choosing these parameters. This allows for the same level of statistical power in resolving anisotropy, while still allowing the spatial resolution to vary. Larger aspect ratios may lead to similar results for the regions in which there are plentiful events. Given the variably-sized regions in Figure 3, the smallest regions will have the lowest maximum values of r. In an effort to ensure uniformity across the regions, we chose a maximum value of $r = 2$ km, which results in a value of $t = 200$ m. We then decreased r by powers of 2, which results in $(r, t) = (1000 \text{ m}, 100 \text{ m}), (500 \text{ m}, 50 \text{ m})$. The latter of these parameter pairs is essentially the lower limit of what is possible, and still have enough points to resolve K_{cyl}.

Since K_{cyl} is a cumulative function of r and t, there may be questions relating to the ability for it to resolve different dip values if present at strictly different length scales. Indeed using such cumulative descriptive metrics is not ideal for this case; a more suitable quantity for this scenario may be the anisotropic pair correlation function, (Møller and Toftaker, 2014; Ross et al., 2022). However, K_{cyl} can still be of some use, depending on the circumstances. To show this, we create a simple synthetic catalog consisting of vertical and horizontal faults having the same strike, as in Figure 5. Here, the vertically dipping faults have an effective length scale of 3 km whereas the horizontal faults have a length scale of 1 km. We compute K_{cyl} for this dataset using $t = 0.25$ km and two values of r, $r = 1$ km, $r = 3$ km. A bootstrap analysis is used to show the dip uncertainty estimates for each value of r. Indeed both faults are reliably recovered.

3 Results

In this section we summarize the main findings for each region and evaluate them in the context of information available from other sources and methods. For southern California, the most comprehensive resource available documenting fault zones and their geometry is the Community Fault Model (CFM) produced by the Southern California Earthquake Center (SCEC; Plesch et al., 2007). This database has been assembled by the SCEC community from a multitude of data sources including focal mechanisms, seismicity, seismic data, geology, and geodetic deformation. The CFM has comprehensive coverage across southern California, and we use version 5.3 (Plesch et al., 2020a) as a baseline for evaluating our results. In addition, we compare our results to those of other studies whenever available, on a case-by-case basis. Next, we walk through the results for each

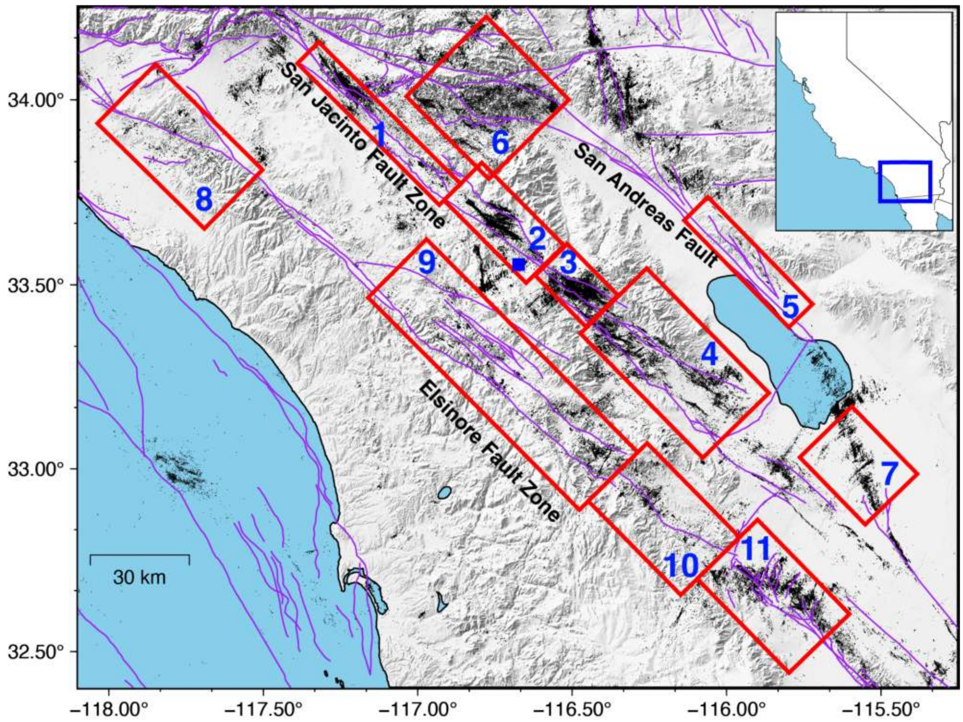


Figure 3 Map of seismicity in Southern California. Black dots indicate relocated epicenters. Red lines denote focus areas with numbers matching region names provided in Table 1. Blue square indicates the town of Anza, California.

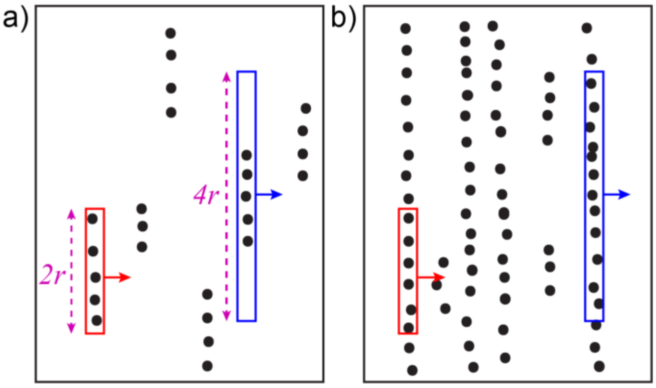


Figure 4 Cartoon illustrating the spatial resolution of the method. In a), the point pattern has an effective length scale of less than $2r$, and the pattern can be resolved by K_{cyl} to $\leq 2r$. In b), the pattern has an effective length scale generally larger than $4r$, but with the two discs shown, the pattern can only be resolved to $\leq 4r$.

fault zone.

3.1 San Jacinto Fault Zone

The San Jacinto Fault Zone (SJFZ) is a major strike-slip system in the southern California plate boundary area that branches off from the San Andreas in the Cajon Pass and extends southeast to the Imperial Valley. The SJFZ has multiple primary strands and several major stepovers (Sharp, 1967). Northwest of the town of Anza, the Clark fault is believed to be the main seismogenic structure of the SJFZ (Share et al., 2017), whereas just southeast of Anza, the Coyote Creek fault branches off of the Clark fault and takes over as the primary fault (Qiu et al., 2017). The seismicity in the SJFZ tends to exhibit weak spatial clustering but strong geometric anisotropy (Ross et al., 2022). The SJFZ exhibits considerable variation in the seismogenic depth along-strike that is attributed to variations in heat flow (Doser and Kanamori, 1986), with depths approaching 20 km at the northwest end in the Cajon Pass, to roughly 10 km near the Salton Trough. While historically considered to be a nearly vertical fault zone, more recent works have concluded that the main structures in the central SJFZ are dipping to the northeast, particularly at depth (Plesch et al., 2020a; Ross et al., 2017a; Schulte-Pelkum et al., 2020). Schulte-Pelkum et al. (2020) conclude that most of the central SJFZ is dipping NE in the range $\sim 65° - 80°$.

We analyze four key seismicity regions of the SJFZ in Figure 6 (see also Table 1) with cylindrical K-functions: Claremont, Hot Springs, Trifurcation area, and Borrego Mountain. The results in Figure 6 are computed over the entire $[0, 22]$ km depth range, and should therefore be interpreted as average values; however it should be noted that for the SJFZ, seismicity generally does not occur above 5 km or so (Hauksson and Meier, 2019), and thus the results largely reflect the deeper part of the fault zone. Each row uses a different combination of (t, r). We notice from the diagrams that in each case, the largest value of K_{cyl} indicates a fault normal azimuth in the range of $29° - 64°$. In fact, except for the Claremont section, the SJFZ regions have a consistent estimate of the fault normal azimuth in the range $29° - 39°$. The radius of the polar plot indicates the dip of the normal vector, and can be used to estimate the average dip of the fault zone; the bootstrap histograms in the bottom row of Figure 6 show the estimated dips and their uncertainties. In the Hot Springs section, $\delta = 68° - 72°$ NE, the Trifurcation area estimates are $\delta = 77° - 84°$ NE, and the Borrego Mountain estimates are $\delta = 75° - 79°$ NE. The

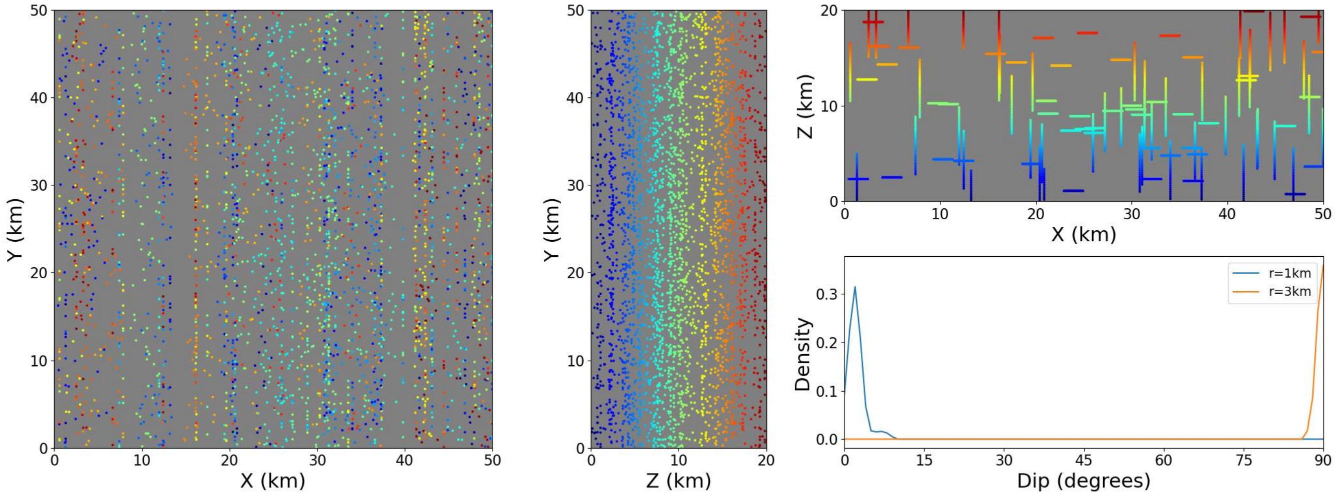


Figure 5 Synthetic catalog demonstration of two fault dip orientations at different length scales (1 km and 3 km, respectively). Events are colored by depth to enhance visibility. Lower right panel shows bootstrap recovery results for K_{cyl} at two different length scales.

SCEC CFM has most of these faults listed as subvertical NE dipping faults, with the Hot Springs, Trifurcation, and Borrego Mountain dip values given as $\delta = 82°$ NE, $\delta = 88° - 89°$ NE, and $\delta = 88° - 89°$ NE, respectively. However, our results for the Claremont section indicate the opposite sense of dip, with δ estimated to be $78° - 88°$ SW; this is in fact close to the CFM results, which has $\delta = 84°$ SW. The results in this figure have effective length scales of 1, 2, and 4 km, and since there is little variation in the dip for these different parameters, they indicate that the dip estimates are robust at these scales. The results do not imply anything about dip at larger scales.

The abundance of seismicity in the central SJFZ allows us to further quantify the dip in depth slices to look for possible depth-dependent variations. Ross et al. (2017a) argued the SJFZ trifurcation area exhibits listric-type behavior based on combined examination of relocated seismicity, focal mechanisms, and mapped surface fault traces. Ross et al. (2017a) concluded that the SJFZ is nearly vertical in the upper 10 km and dipping $70°$ NE below this. Here, we independently investigate this idea with K_{cyl} by splitting the seismicity into three depth bins: 0–8 km, 8–13 km, and >13 km, containing 5584, 16862, and 7466 events, respectively. Figure 7 shows K_{cyl} for the three depth bins. The best estimates of δ are $88°$ NE, $76°$ NE, and $53°$ NE, respectively, which indeed suggest that the fault zone is listric in this area, consistent with the conclusions of Ross et al. (2017a). For cross sections of the seismicity in this area, the reader is recommended to see Figure 7 of Schulte-Pelkum et al. (2020) or Figure 2 of Ross et al. (2017a).

3.2 San Andreas Fault Zone

The portion of the San Andreas Fault Zone (SAFZ) from the Cajon Pass to its terminus at Bombay Beach is just one of the three major sub-parallel strike-slip systems in southern California. There are important questions about its geometry along this part of the plate boundary and it has been the subject of extensive analysis (Fuis et al., 2012, 2017; Lindsey and Fialko, 2013; Fattaruso et al., 2014; Schulte-Pelkum et al., 2020), much of which has focused on whether the main seismogenic fault is vertical or dipping northeast, a question that is of prime importance for earthquake rupture simulations as it will affect both the magnitude of potential earthquakes and also the shaking pattern (Graves et al., 2008, 2011).

The San Gorgonio Pass (SGP) region of the SAFZ is concentrated around the San Bernardino Mountains. The seismicity here is weakly clustered spatially (Ross et al., 2022) and extends down to a depth of $\sim$20 km, the effective lower limit for seismicity in southern California (Hauksson et al., 2012). The slip rate in this area is about 24 mm/year and there are several major strands: the Mission Creek, Banning, and Garnet Hill faults (Gold et al., 2015; Fuis et al., 2017; Blisniuk et al., 2021). There are also numerous minor strands that may not extend to the surface (Fuis et al., 2017; Schulte-Pelkum et al., 2020). Since the start of the instrumental era of seismology in southern California, two significant earthquakes occurred in this area, 1948 M_L 6.5 Desert Hot Springs (Richter et al., 1958; Nicholson, 1996) and 1986 M_w 6.0 North Palm Springs (Jones et al., 1986; Nicholson, 1996; Mori and Frankel, 1990).

Figure 8 shows K_{cyl} results for the SGP region. The estimates of φ and δ indicate a NE dipping fault zone, with δ in the range $54° - 70°$, depending on the scale of the cylindrical elements used. More specifically, we find that δ decreases as the length scale is increased, which suggests that the larger (older) structures in this fault zone are oriented more horizontally, whereas the younger (smaller) structures are slightly more vertical. For comparison, the CFM (Plesch et al., 2020a) has the Banning Fault dipping $72°$ NE and the Mission Creek Fault dipping $82°$ NE. Fuis et al. (2017) identify seismic reflectors in this area that are dipping in the range $\sim 55° - 65°$ NE, with some more steeply dipping structures too. The 1948 M_L 6.5 and 1986 M_w 6.0 main-

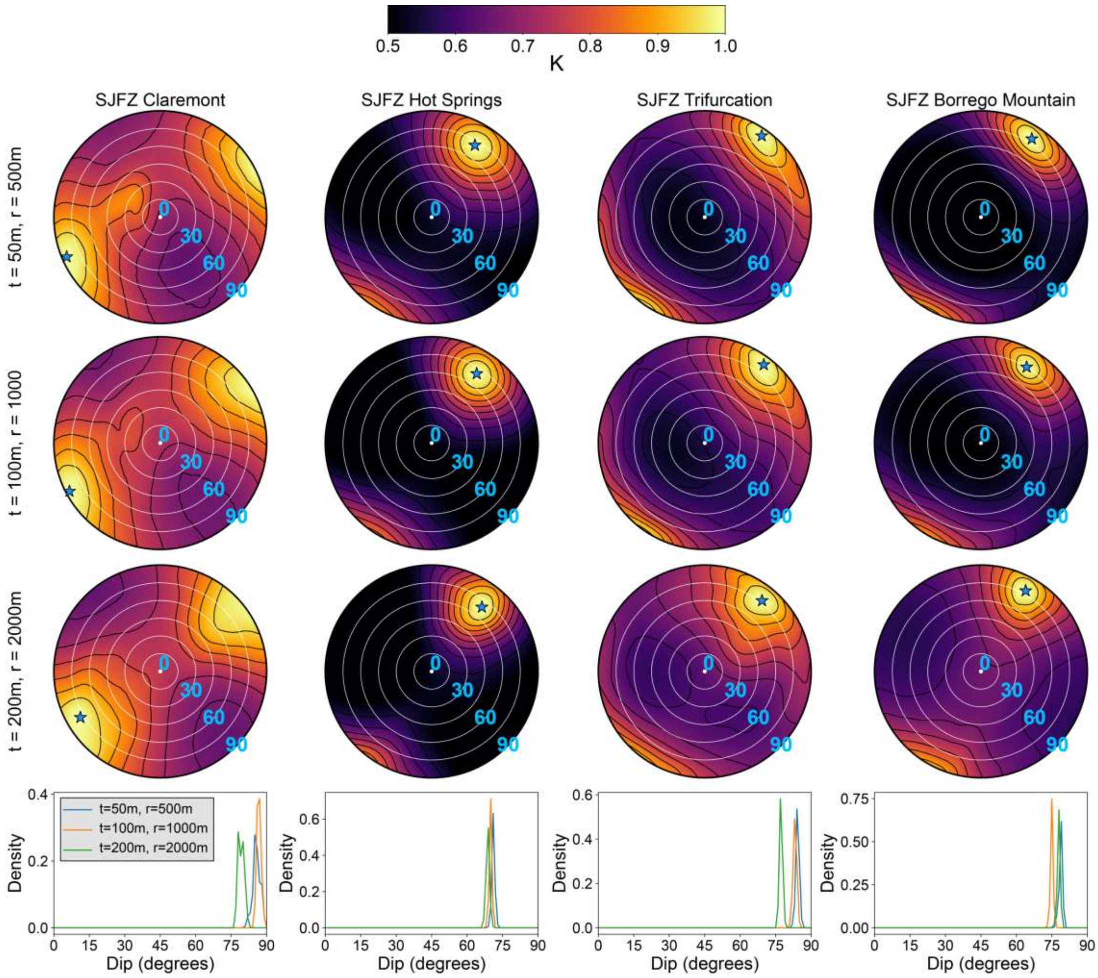


Figure 6 Cylindrical K-functions for the San Jacinto Fault Zone and dip estimates. Density functions for each region (bottom row) are bootstrap distributions for best dip estimate. These areas trend from northwest to southeast. The Claremont section is nearly vertical on average, whereas the other three sections dip moderately to the northeast.

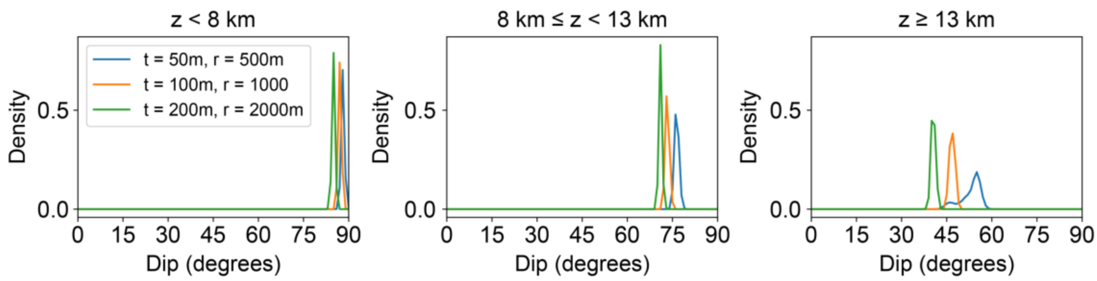


Figure 7 Estimating the depth dependence of δ for the SJFZ Trifurcation Area. This section of the fault zone exhibits evidence of listric strike-slip behavior. Left, middle, and right panels use 5584, 16862, and 7466 events, respectively.

shocks in this area have focal mechanism dips of about $45°$ (Jones et al., 1986; Nicholson, 1996). Our results reflect average values of fault zone dip over the entire SGP region, which includes many smaller structures in between the Banning and Mission Creek faults.

Southeast of the SGP is the Coachella Valley section (Southern San Andreas) of the SAFZ. This portion runs from about Palm Springs to Bombay Beach, the southernmost terminus of the system. In this section also, there is debate over whether the fault zone is dipping (Fuis et al., 2017; Lin et al., 2007; Schulte-Pelkum et al., 2020). The SCEC CFM 5.3 has the Southern San Andreas fault as being pure vertical ($\delta = 90°$), whereas others including Fuis et al. (2017); Lindsey and Fialko (2013) conclude the SAFZ dips $\sim 50° - 60°$ NE. Our K_{cyl} results for the Coachella section of the SAFZ are shown in Figure 8. The method unambiguously identifies a NE dipping fault zone. At the smallest length scale examined, $r = 50$ m, $t = 500$ m, our best estimate of δ is just under $60°$ NE. However, as the scale increases, so does δ: for $r = 100$ m, $t = 1000$ m, $\delta = 73°$ NE, and for the largest scale, $r = 200$ m, $t = 2000$ m, our best estimate of δ is $80°$.

The trend of δ increasing with scale for the Southern San Andreas is opposite to what was observed for the SGP. We interpret these deviations between the smallest and largest scales to reflect down-dip curvature of the fault zone, with a listric type behavior that is more

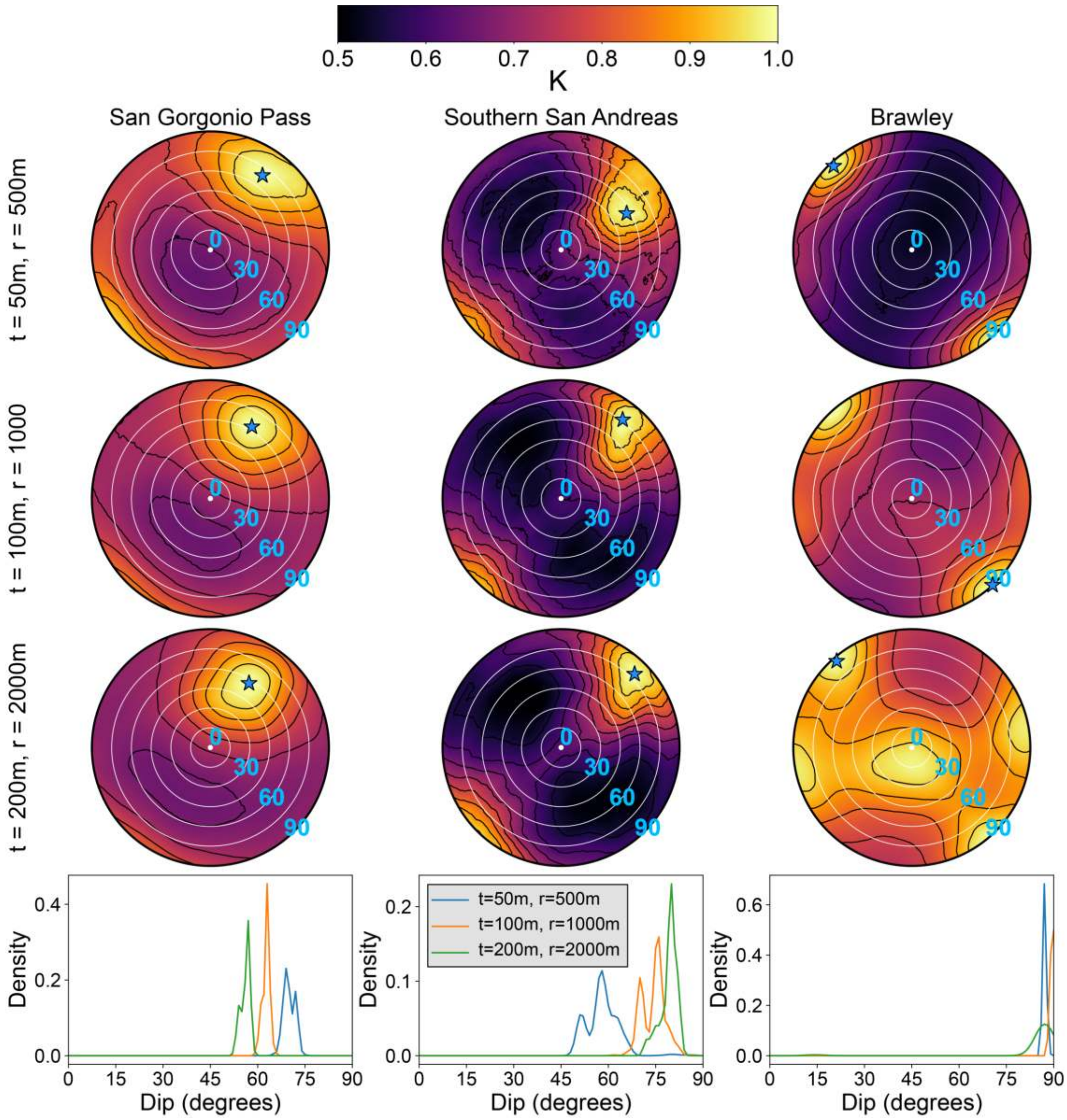


Figure 8 Cylindrical K-functions for the San Andreas Fault Zone and Brawley Seismic Zone. SAFZ seismicity dips to the northeast.

vertical in the upper ~8–10 km and more horizontal below this. However it is important to remember that all scales exhibit clear evidence of a northeast dipping fault zone.

3.3 Brawley Seismic Zone

The Brawley Seismic Zone is one of the more complex faulting regions in California, serving as the plate boundary transition between the SAFZ and the Imperial and Cierro Prieto faults in Baja California. The region is known for having considerable swarm activity (Hauksson et al., 2013, 2017, 2022), conjugate/orthogonal faults (Thatcher and Hill, 1991; Ross et al., 2022), and prolific geothermal activity (Brodsky and Lajoie, 2013).

The SCEC CFM lists all of the major faults in the Brawley Seismic Zone as being vertical strike-slip. Our findings for this region are shown in Figure 8 and have dip estimates that are relatively consistent between the three different length scales. However, there are clear differences in the strike distribution between these scales; the K_{cyl} identifies two clear modes in the strike distribution for $(t = 200\ \mathrm{m},\ r = 2000\ \mathrm{m})$ with roughly equal occurrence, separated by about 60° in azimuth. The conjugate faulting eventually disappears for $(t = 50\ \mathrm{m},\ r = 500\ \mathrm{m})$ and a NW–SE trending orientation is the only one visible. Thus, we can say quantitatively that the NW–SE structures are generally larger than 2 km in length. This orientation is the most closely aligned with the Southern San Andreas, and may reflect the current orientation that new damage and cracking is being produced for. This might imply that the NW–SE trending seismicity structures are relic structure from previous faulting that has not healed.

3.4 Elsinore Fault Zone

The Elsinore Fault Zone (EFZ) is the youngest of the three major fault systems composing the southern California plate boundary. The EFZ also has the lowest slip rates of the three, being ~5 mm/year (Magistrale and Rockwell, 1996). In the northwest, the EFZ emerges near the eastern end of the Los Angeles basin and extends southeast for roughly 200 km before becoming

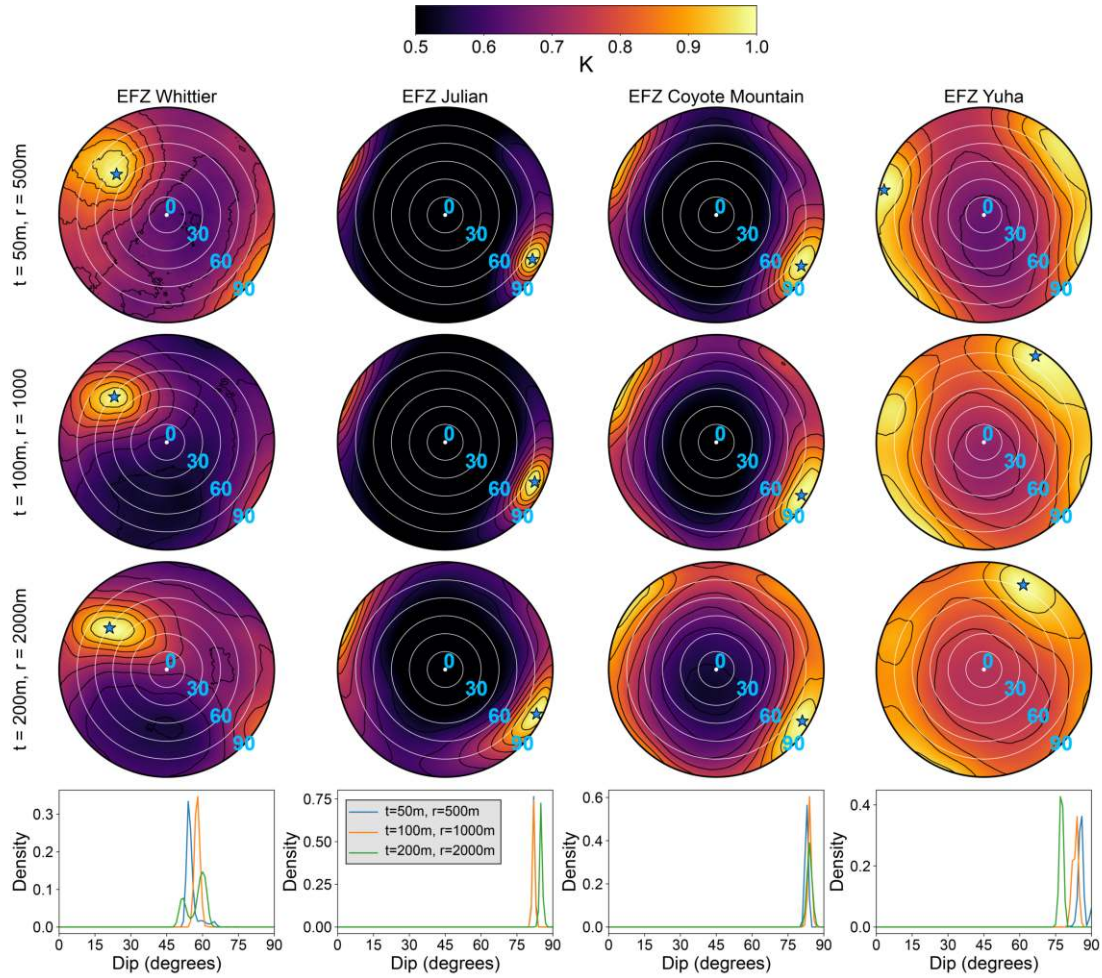


Figure 9 Cylindrical K-functions for the Elsinore Fault Zone. Regions in the first three columns exhibit prominent seismicity anisotropy that is orthogonal to the main strike of EFZ. Most of the EFZ seismicity has nearly vertical dip, except Whittier section. Yuha Desert section has conjugate seismicity with a high angle.

the Laguna Salada Fault Zone near the United States–Mexico border. EFZ seismicity is more scarce compared with some of the other regions we have examined, and so we examine here four sections that have sufficient events to perform a K_{cyl} analysis.

The Whittier section of the EFZ is located in the eastern LA Basin and is viewed as a transition region from the compressional regime of the transverse ranges to the strike-slip regime of the Elsinore system (Hauksson, 1990). The Whittier fault branches off from the dominant trend of the EFZ at an angle of $\sim 15°$ and has a strike of about $300°$. Beneath the Whittier fault is the Puente Hills blind thrust (Shaw and Shearer, 1999). The Whittier fault is listed in the SCEC CFM as dipping to the northeast at $77°$. Events in the area typically have focal mechanisms with considerable obliquity (Yang and Hauksson, 2011), with the largest in recent memory being the 2008 M_w 5.4 Chino Hills earthquake (Hauksson et al., 2008). There has been some discussion of the orientation of the structures here, with both nodal planes being considered as plausible. Shao et al. (2012) analyzed the kinematic rupture process of the Chino Hills earthquake and tested both nodal planes, concluding that the plane aligned with the Whittier fault was most likely. Figure 9 shows our K_{cyl} results for the Whittier, which indicates for all three scales a NW fault zone dipping $51° - 64°$ and a strike of $34° - 40°$. These values are close to the parameters of the "auxiliary plane" for most focal mechanisms in the area; for example the Chino Hills earthquake had an auxiliary plane with a strike of $42°$ and a dip of $55°$. Importantly, K_{cyl} does not show any evidence of a second mode aligned with the Whittier fault. From this, we thus conclude that at least at the scale of 1-4 km, the active seismogenic structures in the area are a mixture of left-lateral and thrust slip that are not aligned with the Whittier fault. At larger scales, it is very possible that fault zone structures align with the Whittier fault and dip northeast as given in the CFM.

The Julian and Coyote Mountain sections cover most of the central and southern EFZ. The major fault traces within these sections are relatively straight and trend southeast. Both sections are listed in the CFM as being nearly vertical ($81° - 87°$) with a strike of around $305°$. For these sections, the peak K_{cyl} value (Figure 9) indicates a strike of $204° - 210°$ and a dip of $82° - 86°$; thus our results identify the orthogonal plane as being the dominant one visible in the seismicity at the scale of 1–4 km. This is similar to the results for the Whittier section. Indeed many of these are large enough to be visible by eye in Figure 3. There is some recognition of the strike direction parallel to the EFZ in the Coyote Mountain results, particularly for the 4 km scale. Therefore the faulting geometry appears to be more complex here and scale dependent.

The final region of the Elsinore that we examine is the Yuha Desert. This area serves as the transition between the Elsinore and Laguna Salada systems and is underlain by the Paso Superior detachment fault (Fletcher et al., 2014). It was the site of extensive aftershock activity following the 2010 M_w 7.2 El Mayor-Cucapah earthquake, including the 2010 M_w 5.7 Ocotillo, California earthquake (Kroll et al., 2013). There also was a shallow M_w 6.5 slow slip event that occurred here as part of this sequence (Ross et al., 2017b). The Yuha Desert area contains numerous fault traces orthogonal to the main trend of EFZ. Indeed our K_{cyl} results corroborate this, with two modes with azimuthal separation of nearly $70°$. At the two largest scales, $t > 100$ m and $r > 1000$ m, the SE trending mode is stronger, whereas, for the smallest scale, the two modes are about equal in strength. There is no evidence for any significant deviation from vertical here, with the $r = 2000$ m scale having a best estimate of $\delta = 78°$ NE.

4 Discussion and conclusions

In this study we have outlined a new method for quantifying the average dip of fault zones using seismicity. Overall our results for southern California seismicity regions compare favorably with those of the SCEC CFM and other sources. While it is just one type of information, it is independent from that considered in the CFM. This study demonstrates the potential for using this method to augment existing CFM databases and ultimately improve upon the known geometrical properties of fault zones.

Our primary findings for the major fault zones examined support the idea that the San Andreas and San Jacinto fault zones in southern California are dipping (at least in an average sense) toward the northeast. Most of the Elsinore Fault Zone is close to vertical, with the lone exception perhaps being the Whittier section at the northwest terminus of the fault zone near the LA Basin. Our findings suggest a progressive steepening of dip spatially, going from SAFZ in the northeast to EFZ in the southwest, which may provide clues as to the tectonic origins of this geometry. These conclusions are consistent with those of Schulte-Pelkum et al. (2020).

Our findings explicitly quantify anisotropy in seismicity at each length scale desired. There are hints of some changes with increasing length scale that may have broader implications about the tectonic history of the region. For example we found minor changes in dip with length scale that may suggest younger faults being formed in recent years may be inconsistent with the larger scale plate boundary faults surrounding them. Additional more detailed analysis is warranted for these cases to further substantiate these observations and possible implications.

The method is not without limitations and these should be emphasized for further clarity on its usage. First, it should be understood that the cylindrical K-function represents average properties over the window. Within the window, the properties may vary spatially, i.e., the seismicity may be viewed as an inhomogeneous point process. While the cylindrical K-function is formulated under the assumption of stationarity, it can still provide useful information even if there are relatively mild deviations from this assumption. An important consequence of the lack of stationarity is that the results will depend on the spatial window chosen. They should be interpreted only for the specific region. This further implies that the results should not be extrapolated to regions outside of the spatial window. Another important limitation results from the "disc" geometry used to construct the cylindrical K-function, which was chosen expressly with the purpose of detecting persistent planar features in seismicity. While not the focus of this study, other types of seismicity features, e.g., linear features, may not be detected with a disc geometry and would require alternatives.

Location errors are the main source of measurement uncertainty in our calculations and their effects should be appropriately considered. The length scales of importance in our study are the values of $2r$, i.e., the diameter of the disc used in computing K_{cyl}. The values used are 1 km, 2 km, and 4 km. The seismicity catalog only included events with successful double-difference relocations and therefore the relative location error is the most important term to consider. For this catalog, 90% of the events are estimated to have relative horizontal and vertical errors of 0.1 km, which is at least an order of magnitude smaller than the length scales considered. We therefore do not expect artifacts related to location errors.

Acknowledgements

The author is grateful to the David and Lucile Packard Foundation for supporting this study through a Packard Fellowship.

Data and code availability

The seismicity catalog used in this study is from Hauksson et al. (2012) and is publicly available from the Southern California Earthquake Data Center (https://scedc.caltech.edu; Southern California Seismic Network, 2013). Maps were created with PyGMT (Uieda et al., 2023).

Competing interests

The authors have no competing interests.

References

Bangs, N. L., Morgan, J. K., Bell, R. E., Han, S., Arai, R., Kodaira, S., Gase, A. C., Wu, X., Davy, R., Frahm, L., Tilley, H. L., Barker, D. H. N., Edwards, J. H., Tobin, H. J., Reston, T. J., Henrys, S. A., Moore, G. F., Bassett, D., Kellett, R., Stucker, V., and Fry, B. Slow slip along the Hikurangi margin linked to fluid-rich sediments trailing subducting seamounts. *Nature Geoscience*, 16(6): 505–512, June 2023. doi: 10.1038/s41561-023-01186-3.

Ben-Zion, Y. and Sammis, C. G. Characterization of Fault Zones. *pure and applied geophysics*, 160(3):677–715, Mar. 2003. doi: 10.1007/PL00012554.

Bird, P. An updated digital model of plate boundaries. *Geochemistry, Geophysics, Geosystems*, 4(3), 2003. doi: 10.1029/2001GC000252.

Blisniuk, K., Scharer, K., Sharp, W. D., Burgmann, R., Amos, C., and Rymer, M. A revised position for the primary strand of the Pleistocene-Holocene San Andreas fault in southern California. *Science Advances*, 7(13):eaaz5691, 2021. doi: 10.1126/sciadv.aaz5691.

Brodsky, E. E. and Lajoie, L. J. Anthropogenic Seismicity Rates and Operational Parameters at the Salton Sea Geothermal Field. *Science*, 341(6145):543–546, Aug. 2013. doi: 10.1126/science.1239213.

Chiaraluce, L., Di Stefano, R., Tinti, E., Scognamiglio, L., Michele, M., Casarotti, E., Cattaneo, M., De Gori, P., Chiarabba, C., Monachesi, G., Lombardi, A., Valoroso, L., Latorre, D., and Marzorati, S. The 2016 Central Italy Seismic Sequence: A First Look at the Mainshocks, Aftershocks, and Source Models. *Seismological Research Letters*, 88(3):757–771, Mar. 2017. doi: 10.1785/0220160221.

Cox, S. F. Injection-Driven Swarm Seismicity and Permeability Enhancement: Implications for the Dynamics of Hydrothermal Ore Systems in High Fluid-Flux, Overpressured Faulting Regimes—An Invited Paper. *Economic Geology*, 111(3):559–587, May 2016. doi: 10.2113/econgeo.111.3.559.

Daley, D. J. and Vere-Jones, D. *An introduction to the theory of point processes: volume I: elementary theory and methods*. Springer, 2003. doi: 10.1007/b97277.

Doser, D. I. and Kanamori, H. Depth of seismicity in the Imperial Valley Region (1977–1983) and its relationship to heat flow, crustal structure and the October 15, 1979, earthquake. *Journal of Geophysical Research: Solid Earth*, 91(B1):675–688, 1986. doi: 10.1029/JB091iB01p00675.

Fattaruso, L. A., Cooke, M. L., and Dorsey, R. J. Sensitivity of uplift patterns to dip of the San Andreas fault in the Coachella Valley, California. *Geosphere*, 10(6):1235–1246, Dec. 2014. doi: 10.1130/GES01050.1.

Fletcher, J. M., Teran, O. J., Rockwell, T. K., Oskin, M. E., Hudnut, K. W., Mueller, K. J., Spelz, R. M., Akciz, S. O., Masana, E., Faneros, G., Fielding, E. J., Leprince, S., Morelan, A. E., Stock, J., Lynch, D. K., Elliott, A. J., Gold, P., Liu-Zeng, J., González-Ortega, A., Hinojosa-Corona, A., and González-García, J. Assembly of a large earthquake from a complex fault system: Surface rupture kinematics of the 4 April 2010 El Mayor–Cucapah (Mexico) Mw 7.2 earthquake. *Geosphere*, 10(4):797–827, Aug. 2014. doi: 10.1130/GES00933.1.

Fuis, G. S., Scheirer, D. S., Langenheim, V. E., and Kohler, M. D. A New Perspective on the Geometry of the San Andreas Fault in Southern California and Its Relationship to Lithospheric Structure. *Bulletin of the Seismological Society of America*, 102(1):236–251, Feb. 2012. doi: 10.1785/0120110041.

Fuis, G. S., Bauer, K., Goldman, M. R., Ryberg, T., Langenheim, V. E., Scheirer, D. S., Rymer, M. J., Stock, J. M., Hole, J. A., Catchings, R. D., Graves, R. W., and Aagaard, B. Subsurface Geometry of the San Andreas Fault in Southern California: Results from the Salton Seismic Imaging Project (SSIP) and Strong Ground Motion Expectations. *Bulletin of the Seismological Society of America*, 107(4):1642–1662, July 2017. doi: 10.1785/0120160309.

Gillard, D., Rubin, A. M., and Okubo, P. Highly concentrated seismicity caused by deformation of Kilauea's deep magma system. *Nature*, 384(6607):343–346, Nov. 1996. doi: 10.1038/384343a0.

Gold, P. O., Behr, W. M., Rood, D., Sharp, W. D., Rockwell, T. K., Kendrick, K., and Salin, A. Holocene geologic slip rate for the Banning strand of the southern San Andreas Fault, southern California. *Journal of Geophysical Research: Solid Earth*, 120(8):5639–5663, 2015. doi: 10.1002/2015JB012004.

Graves, R. W., Aagaard, B. T., Hudnut, K. W., Star, L. M., Stewart, J. P., and Jordan, T. H. Broadband simulations for Mw 7.8 southern San Andreas earthquakes: Ground motion sensitivity to rupture speed. *Geophysical Research Letters*, 35(22), 2008. doi: 10.1029/2008GL035750.

Graves, R. W., Aagaard, B. T., and Hudnut, K. W. The ShakeOut Earthquake Source and Ground Motion Simulations. *Earthquake Spectra*, 27(2):273–291, May 2011. doi: 10.1193/1.3570677.

Hauksson, E. Earthquakes, faulting, and stress in the Los Angeles Basin. *Journal of Geophysical Research: Solid Earth*, 95(B10):15365–15394, 1990. doi: 10.1029/JB095iB10p15365.

Hauksson, E. and Meier, M.-A. Applying Depth Distribution of Seismicity to Determine Thermo-Mechanical Properties of the Seismogenic Crust in Southern California: Comparing Lithotectonic Blocks. *Pure and Applied Geophysics*, 176(3):1061–1081, Mar. 2019. doi: 10.1007/s00024-018-1981-z.

Hauksson, E., Felzer, K., Given, D., Giveon, M., Hough, S., Hutton, K., Kanamori, H., Sevilgen, V., Wei, S., and Yong, A. Preliminary Report on the 29 July 2008 Mw 5.4 Chino Hills, Eastern Los Angeles Basin, California, Earthquake Sequence. *Seismological Research Letters*, 79(6):855–866, Nov. 2008. doi: 10.1785/gssrl.79.6.855.

Hauksson, E., Yang, W. Z., and Shearer, P. M. Waveform Relocated Earthquake Catalog for Southern California (1981 to June 2011). *Bulletin of the Seismological Society of America*, 102(5):2239–2244, Oct. 2012. doi: 10.1785/0120120010.

Hauksson, E., Stock, J., Bilham, R., Boese, M., Chen, X., Fielding, E. J., Galetzka, J., Hudnut, K. W., Hutton, K., Jones, L. M., Kanamori, H., Shearer, P. M., Steidl, J., Treiman, J., Wei, S., and Yang, W. Report on the August 2012 Brawley Earthquake Swarm in Imperial Valley, Southern California. *Seismological Research Letters*, 84(2):177–189, Mar. 2013. doi: 10.1785/0220120169.

Hauksson, E., Meier, M., Ross, Z. E., and Jones, L. M. Evolution of seismicity near the southernmost terminus of the San Andreas Fault: Implications of recent earthquake clusters for earthquake risk in southern California. *Geophysical Research Letters*, 44(3):1293–1301, 2017. doi: 10.1002/2016GL072026.

Hauksson, E., Stock, J. M., and Husker, A. L. Seismicity in a weak crust: the transtensional tectonics of the Brawley Seismic Zone section of the Pacific–North America Plate Boundary in Southern California, USA. *Geophysical Journal International*, 231(1):717–735, Oct. 2022. doi: 10.1093/gji/ggac205.

Hayes, G. P., Wald, D. J., and Johnson, R. L. Slab1.0: A three-dimensional model of global subduction zone geometries. *Journal of Geophysical Research: Solid Earth*, 117(B1), 2012. doi: 10.1029/2011JB008524.

Hayes, G. P., Moore, G. L., Portner, D. E., Hearne, M., Flamme, H., Furtney, M., and Smoczyk, G. M. Slab2, a comprehensive subduction zone geometry model. *Science*, 362(6410):58–61, Oct. 2018. doi: 10.1126/science.aat4723.

Jones, L. M., Hutton, L. K., Given, D. D., and Allen, C. R. The North Palm Springs, California, earthquake sequence of July 1986. *Bulletin of the Seismological Society of America*, 76(6):1830–1837, Dec. 1986. doi: 10.1785/BSSA0760061830.

Kagan, Y. Y. and Knopoff, L. Spatial distribution of earthquakes: the two-point correlation function. *Geophysical Journal International*, 62(2):303–320, Aug. 1980. doi: 10.1111/j.1365-246X.1980.tb04857.x.

Kroll, K. A., Cochran, E. S., Richards-Dinger, K. B., and Sumy, D. F.

Aftershocks of the 2010 M-w 7.2 El Mayor-Cucapah earthquake reveal complex faulting in the Yuha Desert, California. *Journal of Geophysical Research-Solid Earth*, 118(12):6146–6164, Dec. 2013. doi: 10.1002/2013jb010529.

Lay, V., Buske, S., Townend, J., Kellett, R., Savage, M., Schmitt, D. R., Constantinou, A., Eccles, J. D., Gorman, A. R., Bertram, M., Hall, K., Lawton, D., and Kofman, R. 3D Active Source Seismic Imaging of the Alpine Fault Zone and the Whataroa Glacial Valley in New Zealand. *Journal of Geophysical Research: Solid Earth*, 126(12):e2021JB023013, 2021. doi: 10.1029/2021JB023013.

Lin, G., Shearer, P. M., and Hauksson, E. Applying a three-dimensional velocity model, waveform cross correlation, and cluster analysis to locate southern California seismicity from 1981 to 2005. *Journal of Geophysical Research: Solid Earth*, 112 (B12), 2007. doi: 10.1029/2007JB004986.

Lindsey, E. O. and Fialko, Y. Geodetic slip rates in the southern San Andreas Fault system: Effects of elastic heterogeneity and fault geometry. *Journal of Geophysical Research: Solid Earth*, 118(2): 689–697, 2013. doi: 10.1029/2012JB009358.

Loh, J. M. A valid and fast spatial bootstrap for correlation functions. *The Astrophysical Journal*, 681(1):726, 2008. doi: 10.1086/588631.

Magistrale, H. and Rockwell, T. The central and southern Elsinore fault zone, southern California. *Bulletin of the Seismological Society of America*, 86(6):1793–1803, Dec. 1996. doi: 10.1785/BSSA0860061793.

Melgar, D., LeVeque, R. J., Dreger, D. S., and Allen, R. M. Kinematic rupture scenarios and synthetic displacement data: An example application to the Cascadia subduction zone. *Journal of Geophysical Research: Solid Earth*, 121(9):6658–6674, 2016. doi: 10.1002/2016JB013314.

Mori, J. and Frankel, A. Source parameters for small events associated with the 1986 North Palm Springs, California, earthquake determined using empirical Green functions. *Bulletin of the Seismological Society of America*, 80(2):278–295, Apr. 1990. doi: 10.1785/BSSA0800020278.

Møller, J. and Toftaker, H. Geometric Anisotropic Spatial Point Pattern Analysis and Cox Processes. *Scandinavian Journal of Statistics*, 41(2):414–435, 2014. doi: https://doi.org/10.1111/sjos.12041.

Møller, J., Safavimanesh, F., and Rasmussen, J. G. The cylindrical K-function and Poisson line cluster point processes. *Biometrika*, 103(4):937–954, Dec. 2016. doi: 10.1093/biomet/asw044.

Nasirzadeh, F., Shishebor, Z., and Mateu, J. On new families of anisotropic spatial log-Gaussian Cox processes. *Stochastic Environmental Research and Risk Assessment*, 35(2):183–213, Feb. 2021. doi: 10.1007/s00477-020-01906-w.

Nicholson, C. Seismic behavior of the southern San Andreas fault zone in the northern Coachella Valley, California: Comparison of the 1948 and 1986 earthquake sequences. *Bulletin of the Seismological Society of America*, 86(5):1331–1349, Oct. 1996. doi: 10.1785/BSSA0860051331.

Norris, R. J. and Toy, V. G. Continental transforms: A view from the Alpine Fault. *Journal of Structural Geology*, 64:3–31, July 2014. doi: 10.1016/j.jsg.2014.03.003.

Plesch, A., Shaw, J. H., Benson, C., Bryant, W. A., Carena, S., Cooke, M., Dolan, J., Fuis, G., Gath, E., Grant, L., Hauksson, E., Jordan, T., Kamerling, M., Legg, M., Lindvall, S., Magistrale, H., Nicholson, C., Niemi, N., Oskin, M., Perry, S., Planansky, G., Rockwell, T., Shearer, P., Sorlien, C., Suss, M. P., Suppe, J., Treiman, J., and Yeats, R. Community fault model (CFM) for southern California. *Bulletin of the Seismological Society of America*, 97(6): 1793–1802, Dec. 2007. doi: 10.1785/0120050211.

Plesch, A., Marshall, S., Nicholson, C., Shaw, J., Maechling, P., and Su, M. The Community Fault Model version 5.3 and new web-based tools. In *SCEC Annual Meeting, Poster*, volume 184, 2020a. https://www.scec.org/meetings/2020/am/poster/184.

Plesch, A., Shaw, J. H., Ross, Z. E., and Hauksson, E. Detailed 3D Fault Representations for the 2019 Ridgecrest, California, Earthquake Sequence. *Bulletin of the Seismological Society of America*, 110(4):1818–1831, June 2020b. doi: 10.1785/0120200053.

Qiu, H., Ben-Zion, Y., Ross, Z., Share, P.-E., and Vernon, F. Internal structure of the San Jacinto fault zone at Jackass Flat from data recorded by a dense linear array. *Geophysical Journal International*, 209(3):1369–1388, June 2017. doi: 10.1093/gji/ggx096.

Richter, C. F., Allen, C. R., and Nordquist, J. M. The Desert Hot Springs earthquakes and their tectonic environment. *Bulletin of the Seismological Society of America*, 48(4):315–337, Oct. 1958. doi: 10.1785/BSSA0480040315.

Ripley, B. D. The second-order analysis of stationary point processes. *Journal of Applied Probability*, 13(2):255–266, June 1976. doi: 10.2307/3212829.

Rodgers, A. J., Petersson, N. A., Pitarka, A., McCallen, D. B., Sjogreen, B., and Abrahamson, N. Broadband (0–5 Hz) Fully Deterministic 3D Ground-Motion Simulations of a Magnitude 7.0 Hayward Fault Earthquake: Comparison with Empirical Ground-Motion Models and 3D Path and Site Effects from Source Normalized Intensities. *Seismological Research Letters*, 90(3): 1268–1284, May 2019. doi: 10.1785/0220180261.

Ross, Z. E., Hauksson, E., and Ben-Zion, Y. Abundant off-fault seismicity and orthogonal structures in the San Jacinto fault zone. *Science Advances*, 3(3):8, Mar. 2017a. doi: 10.1126/sciadv.1601946.

Ross, Z. E., Rollins, C., Cochran, E. S., Hauksson, E., Avouac, J.-P., and Ben-Zion, Y. Aftershocks driven by afterslip and fluid pressure sweeping through a fault-fracture mesh. *Geophysical Research Letters*, page 2017GL074634, Jan. 2017b. doi: 10.1002/2017GL074634.

Ross, Z. E., Trugman, D. T., Hauksson, E., and Shearer, P. M. Searching for hidden earthquakes in Southern California. *Science*, 364 (6442):767–771, May 2019. doi: 10.1126/science.aaw6888.

Ross, Z. E., Cochran, E. S., Trugman, D. T., and Smith, J. D. 3D fault architecture controls the dynamism of earthquake swarms. *Science*, 368(6497):1357–1361, June 2020. doi: 10.1126/science.abb0779.

Ross, Z. E., Ben-Zion, Y., and Zaliapin, I. Geometrical properties of seismicity in California. *Geophysical Journal International*, 231 (1):493–504, Oct. 2022. doi: 10.1093/gji/ggac189.

Rubin, A. M., Gillard, D., and Got, J.-L. Streaks of microearthquakes along creeping faults. *Nature*, 400(6745):635–641, Aug. 1999. doi: 10.1038/23196.

Safavimanesh, F. and Redenbach, C. A comparison of functional summary statistics to detect anisotropy of three-dimensional point patterns. *arXiv:1604.04211 [stat]*, Apr. 2016. http://arxiv.org/abs/1604.04211.

Sato, H., Hirata, N., Koketsu, K., Okaya, D., Abe, S., Kobayashi, R., Matsubara, M., Iwasaki, T., Ito, T., Ikawa, T., Kawanaka, T., Kasahara, K., and Harder, S. Earthquake Source Fault Beneath Tokyo. *Science*, 309(5733):462–464, July 2005. doi: 10.1126/science.1110489.

Schulte-Pelkum, V., Ross, Z. E., Mueller, K., and Ben-Zion, Y. Tectonic Inheritance With Dipping Faults and Deformation Fabric in the Brittle and Ductile Southern California Crust. *Journal of Geophysical Research: Solid Earth*, 125(8):e2020JB019525, 2020. doi: 10.1029/2020JB019525.

Shao, G., Ji, C., and Hauksson, E. Rupture process and energy budget of the 29 July 2008 Mw 5.4 Chino Hills, California, earthquake. *Journal of Geophysical Research: Solid Earth*, 117(B7), 2012. doi: 10.1029/2011JB008856.

Share, P.-E., Ben-Zion, Y., Ross, Z. E., Qiu, H., and Vernon, F. L. Internal structure of the San Jacinto fault zone at Blackburn Saddle from seismic data of a linear array. *Geophysical Journal International*, 210(2):819–832, Aug. 2017. doi: 10.1093/gji/ggx191.

Sharp, R. V. San Jacinto Fault Zone in the Peninsular Ranges of Southern California. *GSA Bulletin*, 78(6):705–730, June 1967. doi: 10.1130/0016-7606(1967)78[705:SJFZIT]2.0.CO;2.

Shaw, B. E., Milner, K. R., Field, E. H., Richards-Dinger, K., Gilchrist, J. J., Dieterich, J. H., and Jordan, T. H. A physics-based earthquake simulator replicates seismic hazard statistics across California. *Science Advances*, 4(8):eaau0688, Aug. 2018. doi: 10.1126/sciadv.aau0688.

Shaw, J. H. and Shearer, a. P. M. An Elusive Blind-Thrust Fault Beneath Metropolitan Los Angeles. *Science*, 283(5407):1516–1518, Mar. 1999. doi: 10.1126/science.283.5407.1516.

Shearer, P. M. Parallel fault strands at 9-km depth resolved on the Imperial Fault, Southern California. *Geophysical Research Letters*, 29(14):19–1–19–4, 2002. doi: 10.1029/2002GL015302.

Shelly, D. R., Hardebeck, J. L., Ellsworth, W. L., and Hill, D. P. A new strategy for earthquake focal mechanisms using waveform-correlation-derived relative polarities and cluster analysis: Application to the 2014 Long Valley Caldera earthquake swarm. *Journal of Geophysical Research: Solid Earth*, 121 (12):8622–8641, 2016. doi: 10.1002/2016JB013437.

Southern California Seismic Network. Southern California Earthquake Data Center, 2013. doi: 10.7909/C3WD3xH1.

Thatcher, W. and Hill, D. P. Fault orientations in extensional and conjugate strike-slip environments and their implications. *Geology*, 19(11):1116–1120, Nov. 1991. doi: 10.1130/0091-7613(1991)019<1116:FOIEAC>2.3.CO;2.

Uieda, L., Tian, D., Leong, W. J., Schlitzer, W., Grund, M., Jones, M., Fröhlich, Y., Toney, L., Yao, J., Magen, Y., Tong, J.-H., Materna, K., Belem, A., Newton, T., Anant, A., Ziebarth, M., Quinn, J., and Wessel, P. PyGMT: A Python interface for the Generic Mapping Tools. doi: 10.5281/zenodo.7772533.

Yang, W. and Hauksson, E. Evidence for Vertical Partitioning of Strike-Slip and Compressional Tectonics from Seismicity, Focal Mechanisms, and Stress Drops in the East Los Angeles Basin Area, California. *Bulletin of the Seismological Society of America*, 101(3):964–974, June 2011. doi: 10.1785/0120100216.

doi:10.26443/seismica.v3i1.1084

SEISMICA

Testing the Predictive Power of b Value for Italian Seismicity

Cataldo Godano *[1,2], **Anna Tramelli** [2], **Giuseppe Petrillo** [3,4], **Vincenzo Convertito** [2]

[1]Department of Mathematics and Physics, Università della Campania - Luigi Vanvitelli, Caserta, Italy., [2]Istituto Nazionale di Geofisica e Vulcanologia - Sezione di Napoli Osservatorio Vesuviano, Napoli, Italy., [3]The Institute of Statistical Mathematics, Research Organization of Information and Systems, Tokyo 106-8569, Japan., [4]Scuola Superiore Meridionale, Naples, Italy

Author contributions: *Conceptualization* Cataldo Godano, Vincenzo Convertito, Anna Tramelli. *Formal Analysis* Cataldo Godano, Vincenzo Convertito, Anna Tramelli, Giuseppe Petrillo. *Writing - Original draft* Cataldo Godano, Vincenzo Convertito, Anna Tramelli, Giuseppe Petrillo.

Production Editor: Gareth Funning
Handling Editor: Andrea Llenos
Copy & Layout Editor: Anant Hariharan

Received: July 10th, 2023
Accepted: January 16th, 2024
Published: February 1st, 2024

Abstract A very efficient method for estimating the completeness magnitude m_c and the scaling parameter b of earthquake magnitude distributions has been thoroughly tested using synthetic seismic catalogues. Subsequently, the method was employed to assess the capability of the b value in differentiating between foreshocks and aftershocks, confirming previous findings regarding the Amatrice-Norcia earthquake sequence. However, a blind algorithm reveals that the discriminative ability of the b value necessitates a meticulous selection of the catalogue, thereby reducing the predictability of large events occurring subsequent to a prior major earthquake.

1 Introduction

The exponential earthquake magnitude distribution, known as the Gutenberg and Richter (GR) law (Gutenberg and Richter, 1944), establishes that:

$$p(m) = b\ln(10)10^{-bm} \tag{1}$$

The scaling parameter b has been extensively investigated because it represents a primary instrument for the evaluation of the occurrence probability of an earthquake of a given size. Its value has been inversely correlated to the stress state (Scholz, 1968; Wyss, 1973; Amitrano, 2003; Gulia and Wiemer, 2010), attracting research interest on its spatial and temporal variations.

Generally, analyses of spatial variations of the b value are performed by mapping it on a regularly spaced grid. The inclusion of the earthquakes in the cells can follow different rules (minimum number of events, maximum distance from the centre of the cell, etc.) (Wiemer and Wyss, 1997, 2002) producing, in some cases, overlapping cells or, in other cases, earthquakes that are not included in any cell (Kamer and Hiemer, 2015; Godano et al., 2022). This may prevent a formally correct statistical comparison between different cells of the grid. Some authors weight each earthquake on the basis of its distance from the grid node of interest (Tormann et al., 2014). Many authors have applied this method to several regions of the world (see, among others, Kamer and Hiemer (2015); Taroni et al. (2021); García-Hernández et al. (2021); Pino et al. (2022)). However, it introduces correlations in the grid of the b values. Recently, Godano et al. (2022) introduced a parameter-free method producing fully independent b values and reducing the number of missed earthquakes.

Decreases of b have been proposed to indicate the occurrence of foreshocks before a large earthquake (Papadopoulos, 1988; Papadopoulos et al., 2018, 2010) or to characterize the stress field in a volcanic area (Tramelli et al., 2021). In a recent paper, Gulia and Wiemer (2019) suggested that a smaller b can discriminate between foreshocks and mainshocks in seismic sequences. However, Lombardi (2022) questioned this result because the completeness magnitude m_c (see next paragraph for details on this parameter) could be biased, causing a biased estimation of the b value.

Here, we perform a detailed analysis of the method introduced by Godano et al. (2023) for m_c evaluation in order to verify its reliability, and consequently, the reliability of the connected b value estimations. Then, we apply the method to real-time discrimination of earthquake foreshocks and aftershocks for the Amatrice - Norcia sequence in Italy, following the selection procedure adopted by Gulia and Wiemer (2019) and adopting a blind procedure for the earthquake selection.

2 The evaluation of m_c

The magnitude of completeness, m_c, is defined as the lowest magnitude at which earthquakes are reliably recorded and reported in earthquake catalogues (Rydelek and Sacks, 1989). Its evaluation is extremely important because its underestimation will cause an underestimation of the b value. Conversely, its overestimation implies a loss of information and a bias in the deter-

*Corresponding author: cataldo.godano@unicampania.it

mination of the b value, due to the reduction of magnitude range. Several methods have been proposed for estimating the m_c value (Wiemer and Wyss, 2000; Cao and Gao, 2002; Ohmura and Kawamura, 2007; Godano, 2017; Godano et al., 2023; Godano and Petrillo, 2023; Roberts et al., 2015). In the following list we report and discuss some of the methods that are based on earthquake catalogues, even though there exist other methods (not considered here) that are based on seismic networks (Mignan and Woessner, 2012; Tramelli et al., 2013).

- *The maximum curvature technique* (Wiemer and Wyss, 2000) recognizes m_c as the magnitude at which the Gutenberg-Richter law reaches its maximum value. This method tends to underestimate m_c (Mignan and Woessner, 2012), and consequently, also the b value. Indeed, Wiemer (2001) add 0.2 to their estimated m_c, which is, in some sense, arbitrary.

- *The goodness of fit test* (Wiemer and Wyss, 2000) evaluates the correlation coefficient r of the linearized expression of the Gutenberg-Richter law as a function of a threshold magnitude m_{th}. When r reaches its maximum, or a stable value, then $m_{th} = m_c$. The method presents the disadvantage that the r value for a linearized expression could be, in some cases, unstable, leading to a biased estimation of m_c and b (Mignan and Woessner, 2012).

- *The harmonic mean method* (Godano, 2017) is based on the observation that the harmonic mean of an exponential distribution increases linearly with m_{th}. Consequently, it deviates from linearity for $m_{th} < m_c$. Although the method presents some advantages, similar to the goodness of fit test, the instability of r can produce a biased estimation of m_c and b.

- *The entire magnitude range method* (Ogata and Katsura, 1993) multiplies the Gutenberg-Richter relationship in Eqn. (1) by the cumulative Gaussian distribution of the parameters μ and σ. This implies that at $m = \mu + \sigma$, 50% of the earthquakes are recorded in the catalogue and this probability increases to 95% if $m = \mu + 2\sigma$. The advantage of the method is represented by the possibility of defining a magnitude probability density function for the whole range of magnitudes present in the catalogue. However, above the best completeness value, this method does not describe the gradual curvature of the GR distribution correctly by not multiplying the detection function by the theoretical GR law (Mignan and Woessner, 2012).

- *The b value stability approach* (Cao and Gao, 2002) evaluates the b value as a function of m_{th} and considers $m_{th} = m_c$ when b reaches a stable value. The problem with such a method is the strong fluctuations of the b value at higher m_{th} due to undersampling (Mignan and Woessner, 2012).

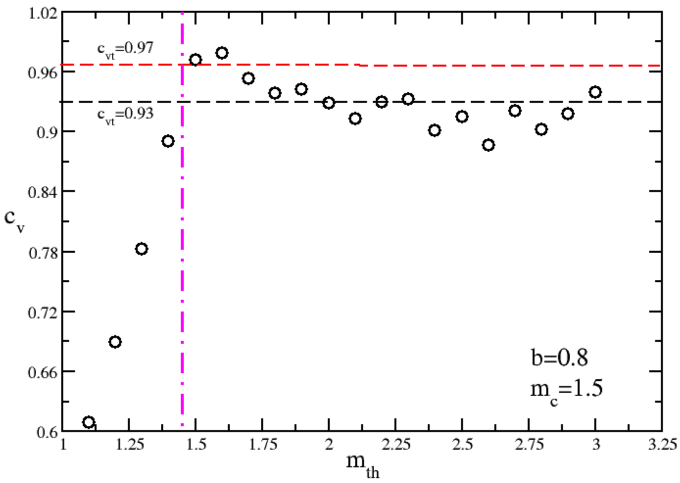


Figure 1 c_v as a function of m_{th} for a simulated catalogue. The dashed lines represent two values of c_{vt}. The vertical dot-dashed magenta line represents point from which we select the best m_{th} The values of b and m_c used to produce the catalogue are also reported in the figure.

2.1 Estimating m_c using the c_v based method.

In the following we will use the method introduced by Godano et al. (2023). The method can be considered a generalization of the one introduced by Cao and Gao (2002) based on b value stabilization. Conversely, Godano et al. (2023) introduce second-order statistics, observing that the variability coefficient (defined as the ratio between the standard deviation and the average value) of an exponential distribution assumes a value equal to 1. More precisely, we define the quantity $m_1 = m - m_{th}$ and evaluate its variability coefficient

$$c_v = \frac{\sigma_{th}}{\langle m_1 \rangle} \tag{2}$$

with σ_{th} being the standard deviation of m_1 and $\langle m_1 \rangle$ its average value. m_1 must follow an exponential distribution whose c_v assumes the value 1. As a consequence, when evaluating c_v as a function of a threshold magnitude m_{th}, c_v assumes a value $\simeq 1$ at $m_{th} = m_c$, where the distribution becomes a purely exponential distribution. An example for a simulated catalogue (see next paragraph for details) is shown in Fig. 1. As can be seen, c_v does not assume the value of 1, typical of a purely exponential distribution, for all the m_{th} values. This occurs because of the fluctuation of the distribution around the purely exponential one. In other cases (not shown here), c_v can assume values slightly larger than 1. For this reason, it is opportune to introduce a c_v threshold value (let us call it c_{vt}) above which the distribution can be considered a purely exponential distribution. More precisely, we choose m_c to be the smallest m_{th} where c_v is larger than the threshold c_{vt}. Here we show, as an example, two values of c_{vt}: 0.93 and 0.97. In both cases the m_c value is correctly identified for the example shown here. A more accurate investigation of the appropriate c_{vt} value is performed below.

Let us test the reliability of the method by means of some simulations.

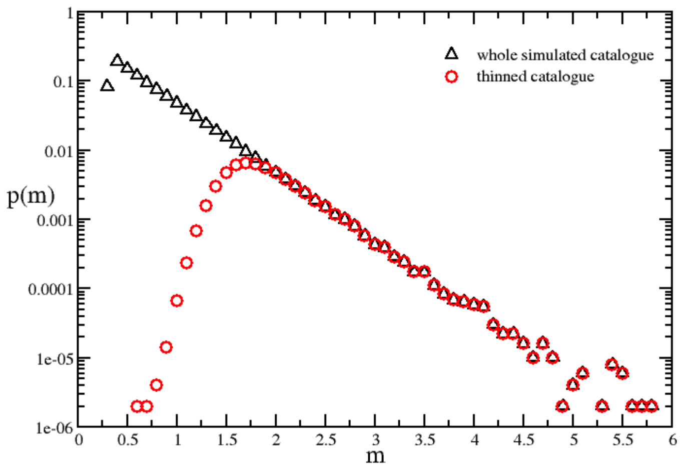


Figure 2 An example of the GR distribution for a simulated catalogue before and after thinning.

2.2 Randomly simulated catalogues

In order to test the reliability of the method, we simulated 10^5 catalogues with 20000 earthquakes and a b value randomly chosen from a uniform distribution in the range of [0.5,1.5]. Then, each catalogue is thinned using the Ogata and Katsura (1993) approach. The method consists of multiplying the Gutenberg-Richter distribution by an $erfc(m)$ function. This provides the number of events to be removed from the catalogue for each $m < m_c$. While simulated catalogues can be truncated at a given threshold magnitude, this does not correctly simulate experimental catalogues containing some, though not all, events with $m < m_c$. The $erfc(m)$ parameters have been selected on the basis of the following rules: μ has been randomly generated in the range [1.5,2.5], whereas σ is fixed at 0.1 considering $m_c = \mu + 2\sigma$. An example of the GR distribution for a simulated catalogue, before and after thinning, is shown in Fig. 2. Then, for each catalogue, we estimate m_c using the Godano et al. (2023) method and b by means of the standard maximum likelihood method (Aki, 1965) and evaluate the quantities $\Delta b = b - b_e$ and $\Delta m_c = m_c - m_{c_e}$, where b and m_c are the parameters used for the simulation and b_e and m_{c_e} are the corresponding estimated values. The distributions of Δb and Δm_c are then evaluated.

There are three quantities affecting Δb and Δm_c, namely N, the minimum number of events in the range $m_{max} - m_c$, the range itself $\Delta m = m_{max} - m_c$, and c_{vt}. More precisely, we evaluate the distributions of Δb and Δm_c using only catalogues:

1. with a number n of events with $m \geq m_c$ larger than or equal to a given value of N without any restriction on Δm and using c_{vt}=0.97
2. with Δm larger than a given value without any restriction on N and using c_{vt}=0.97
3. with different values of c_{vt} without any restriction on Δm and $N = 100$ (this restriction is adopted to evaluate a reliable value of σ_{th})

The evaluated distributions are reported in Figures 3 to 5. In all cases the distributions are sharply peaked (indicating supergaussian distributions) at $\Delta b = 0$ and at $\Delta m_c \simeq -0.25$, revealing a small tendency to overestimate m_c. In the supplementary information we report the results of the Kolmogorov-Smirnov test at a 99% confidence level (test statistic= 0.31) in order to reject the hypothesis that the samples follow a Gaussian distribution. The results are not strongly influenced by N and Δm, although for $N = 300$ and $\Delta m = 3$ long tails in Δb distributions are avoided. However, for these values more than 50% of the catalogues are discarded due to a small number of events or too small a magnitude range.

The overestimation of m_c when c_{vt}=0.97 suggests that this parameter also influences our results. As a consequence we perform a sensitivity analysis, varying c_{vt} in the range [0.93,0.99]. Fig. 5 reveals that $p(\Delta m_c)$ is correctly peaked at $\Delta m_c = 0$ when c_{vt}=0.93, which can be assumed as the best value for c_{vt}.

2.3 Epidemic Type Aftershock Sequence (ETAS) simulated catalogues

The ETAS model represents the gold standard for testing seismic clustering hypotheses and forecasting (Ogata, 1988, 1998; Helmstetter and Sornette, 2003; Console et al., 2007; Lombardi and Marzocchi, 2010; Zhuang, 2011, 2012). In this model the occurrence rate λ of an event with magnitude $m > m_0$, at a position (x, y) and a time t, can be written as

$$\lambda(x, y, t|\mathcal{H}_t) = \mu(x, y)+ \sum_{j:t_j<t} K10^{\alpha(m_j-m_0)}\frac{(p-1)c^{p-1}}{(t-t_j+c)^p}\frac{q-1}{\pi}(\delta(m_j))^{q-1}[(x-x_j)^2 + (y-y_j)^2 + \delta(m_j)]^{-q} \quad (3)$$

where the sum extends over all previous events that occurred in a certain region, $\mu(x, y)$ describes the time-stationary background rate, $\delta(m) = d10^{\gamma m}$ from Kagan (2002) and the set $\hat{\theta} = (p, \alpha, c, K, d, q, \gamma)$ contains the fitting parameters of the model. The estimation of the set $\hat{\theta}$ that best fits the experimental data can be performed using maximum likelihood methods (Lippiello et al., 2014; Ogata, 1998; Y., 1983; Ogata and Zhuang, 2006) or different types of tuning techniques (Petrillo and Lippiello, 2020, 2023).

The generation of an ETAS catalog is a standard procedure described in Zhuang et al. (2004); Zhuang and Touati (2015) and de Arcangelis L. et al. (2016). The first step is setting the background seismicity $\mu(x, y)$. This represents the zeroth order generation in a self-exciting branching process and a certain number n_0 of events are created. Each of these elements generate a certain number of offspring, i.e, the aftershocks. The number n_1, the occurrence, and the spatiotemporal position of the aftershocks depend on the functional form $\lambda(x, y, t)$ and on the parameters $\hat{\theta}$. In practice, the number of aftershocks is extracted from a Poisson distribution with an average dictated by the productivity law. For each offspring, the occurrence time is extracted based on the Omori-Utsu law and the location is based on the spatial

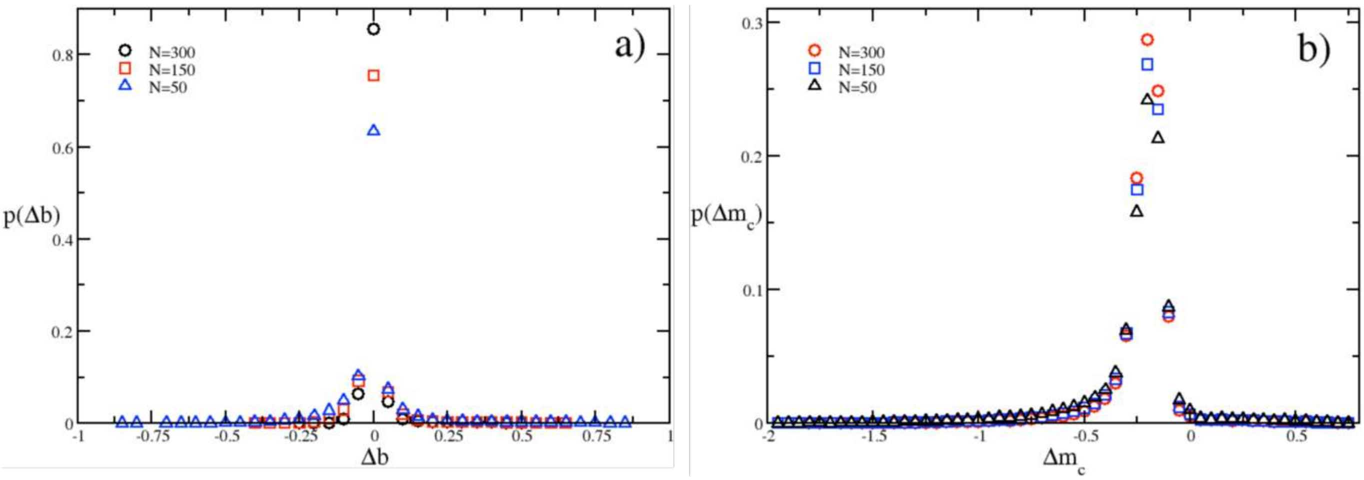


Figure 3 Distributions of Δb (panel (a)) and Δm_c (panel (b)) for different values of N and c_{vt}=0.97.

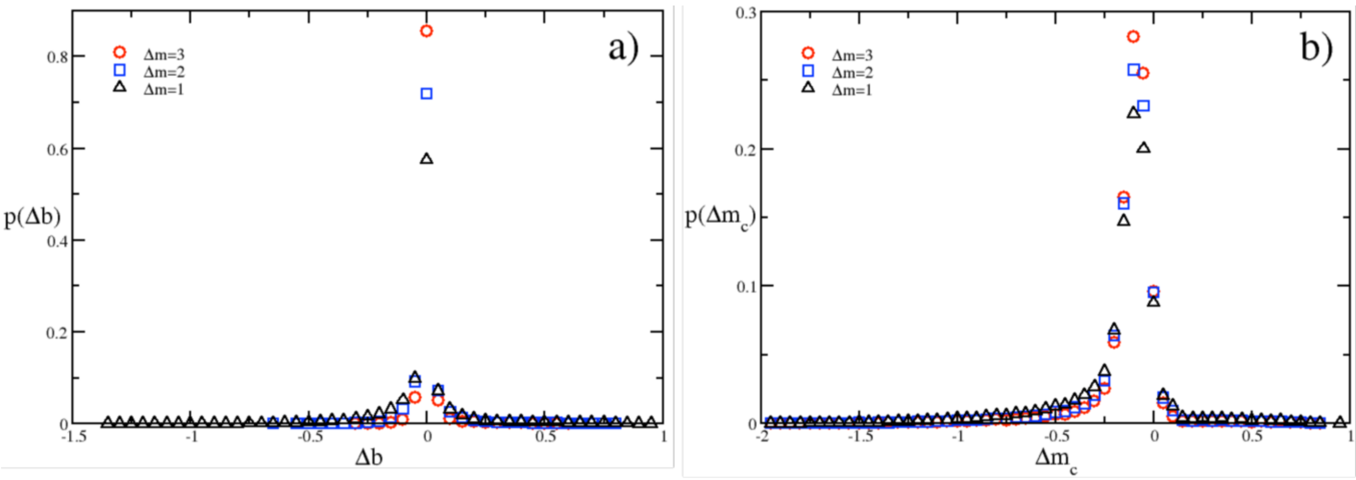


Figure 4 The distributions of Δb (panel (a)) and Δm_c (panel (b)) for different values of Δm and c_{vt}=0.97.

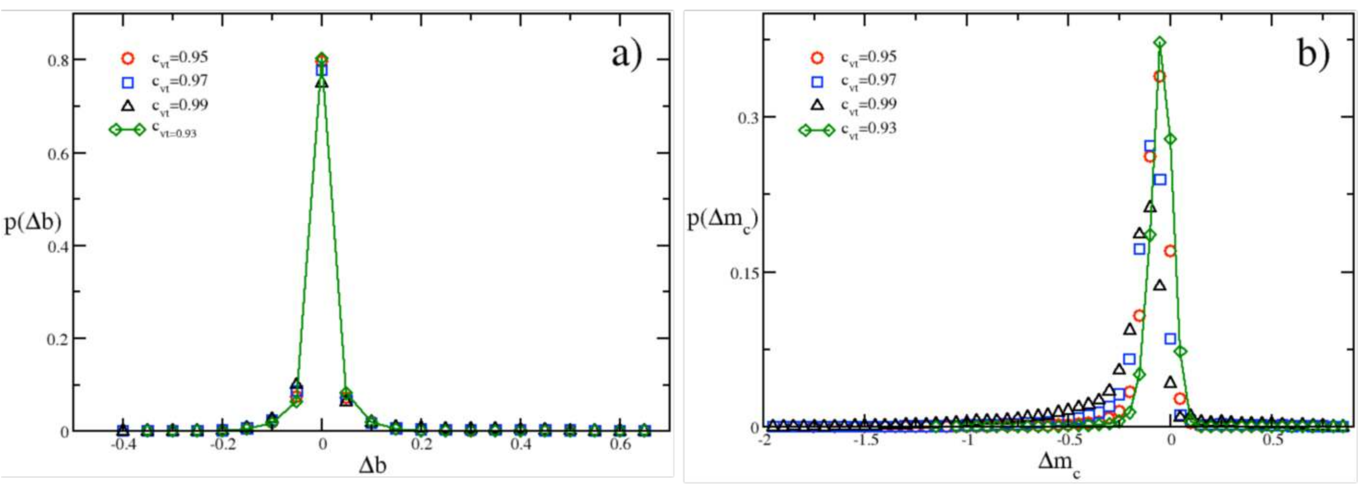


Figure 5 The distributions of Δb and Δm_c for different values of c_{vt}. Here N and Δm are fixed at 100 and 2 respectively.

distribution. As a last step, the magnitude of the event is assigned obtaining the value from the Gutenberg-Richter law in Eqn. (1), since we are assuming magnitude independence among triggered events (Petrillo and Zhuang, 2022, 2023). This is the first-order generation of events. The previous step is repeated considering $n_j = n_{j-1}$ and it is iterated until $n_{j*} = 0$. We would like to emphasize that the numerical catalogs generated using this method do not exclude any events; in other words, no completeness threshold is used to get the simulation. In this study we employ the parameters optimized by Petrillo and Lippiello (2023). In order to verify the reliability of the method in analyzing time variations of the b value we simulated 5 ETAS (Ogata, 1999) catalogues with about 10^6 events and different values of b, in particular $b = 0.6, 0.8, 1.0, 1.2, 1.4$. Then, employ-

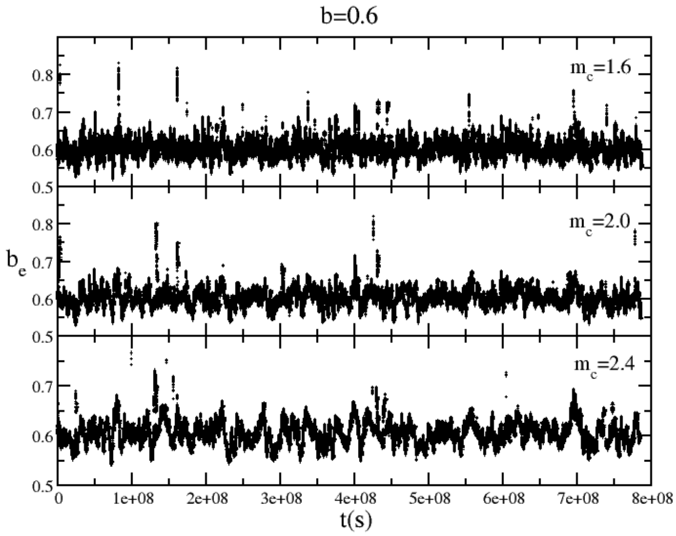


Figure 6 The time variation of the b_e value for the catalogue with b=0.6 thinned at three different m_c values.

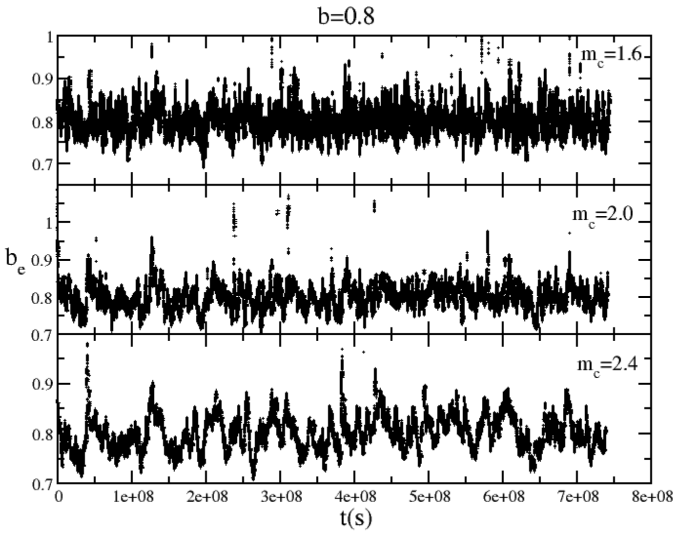


Figure 7 The time variation of the b_e value for the catalogue with b=0.8 thinned at three different m_c values.

ing the same thinning procedure used before, for each value of b we obtain 3 different incomplete catalogues by setting $m_c = 1.6, 2.0,$ and 2.4, for a total of 15 synthetic incomplete catalogues. The temporal variations of the b value are finally obtained by considering windows of 1000 events, sliding on one event at a time. For each window we evaluate m_c with the Godano et al. (2023) method and b by maximizing the likelihood (Aki, 1965). In this case we use $N = 150$, Δm = 2 and $c_{vt} = 0.93$. Of course, we expect no time variation of the b and m_c values or, at least, weak fluctuations of their values around the 'true' values. However Figures 6 - 10 show that b appears to fluctuate around an underestimated value. In general, the changes in the b value could reflect the physical processes of stress evolution and crack growth. However, the fluctuations observed in the analyses are within statistical error, so in this case, a decrease in b is not an indication of precursor phenomena. For high values of the magnitude of completeness, some gaps are present in $b_e(t)$. This is explained by the fact that in data

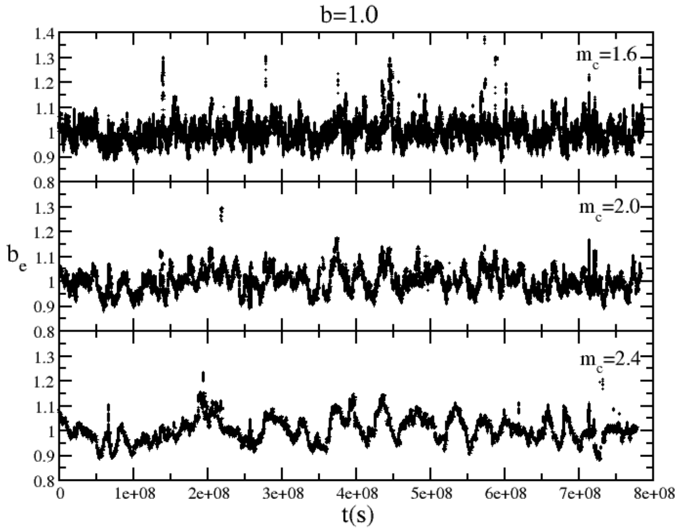


Figure 8 The time variation of the b_e value for the catalogue with b=1.0 thinned at three different m_c values.

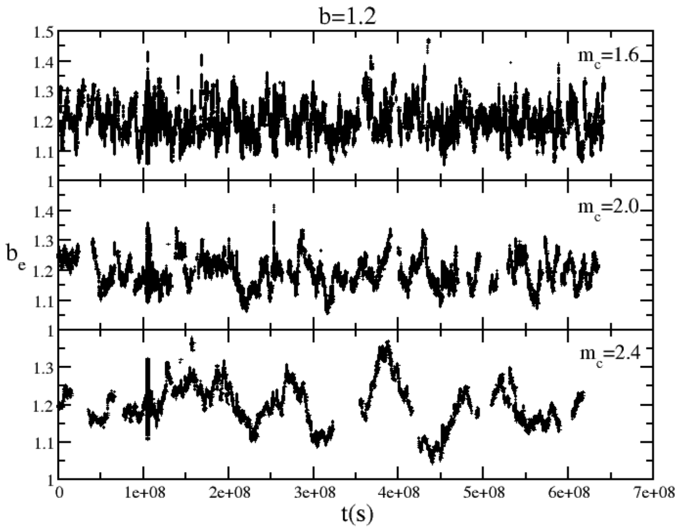


Figure 9 The time variation of the b_e value for the catalogue with b=1.2 thinned at three different m_c values.

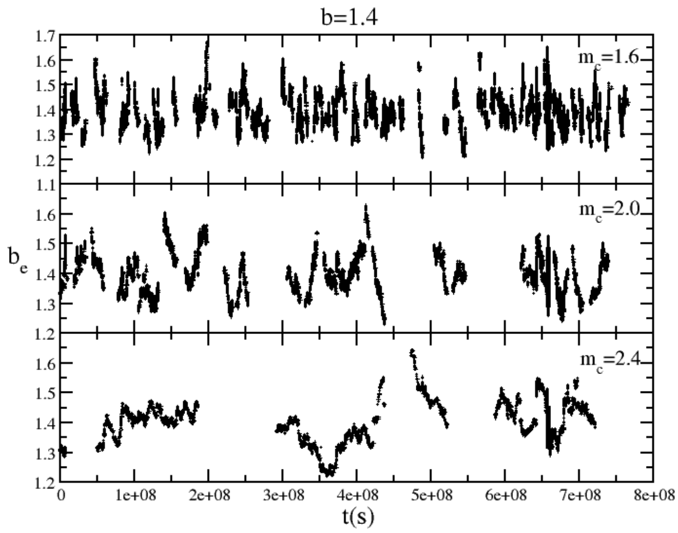


Figure 10 The time variation of the b_e value for the catalogue with b=1.4 thinned at three different m_c values.

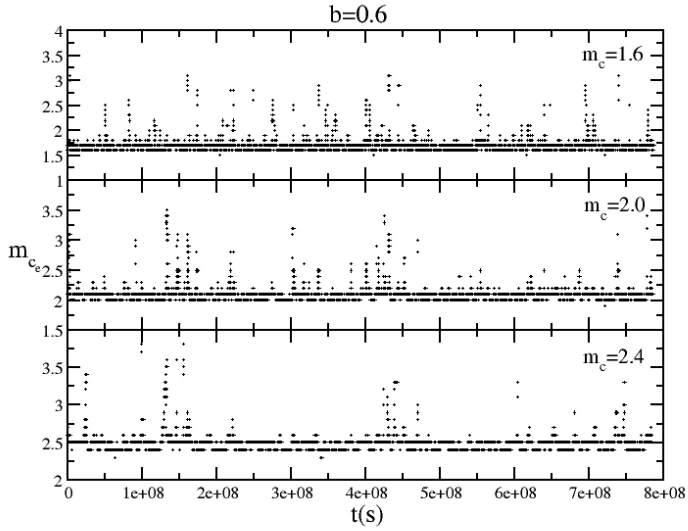


Figure 11 The time variation of the m_{c_e} value for the catalogue with b=0.6 thinned at three different m_c values.

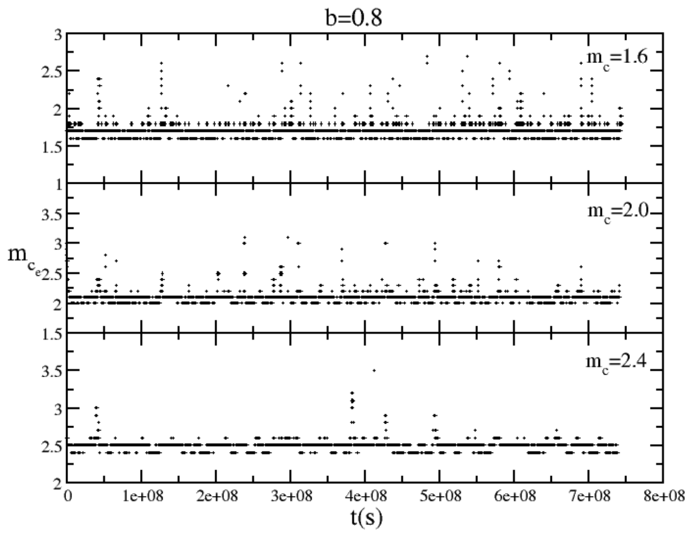


Figure 12 The time variation of the m_{c_e} value for the catalogue with b=0.8 thinned at three different m_c values.

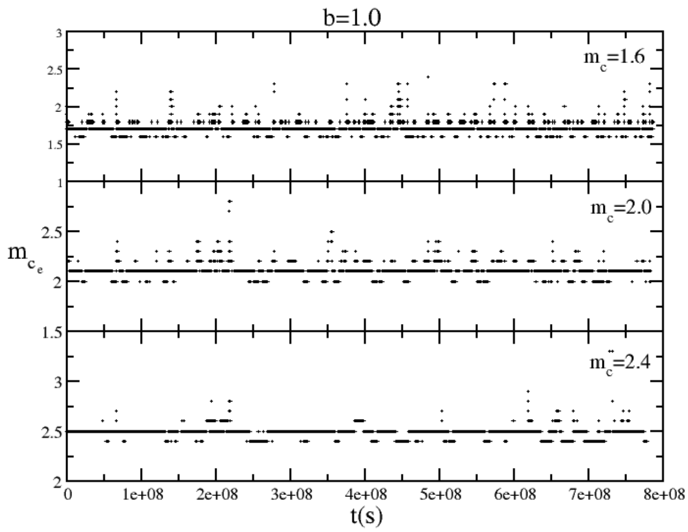


Figure 13 The time variation of the m_{c_e} value for the catalogue with b=1.0, thinned at three different m_c values.

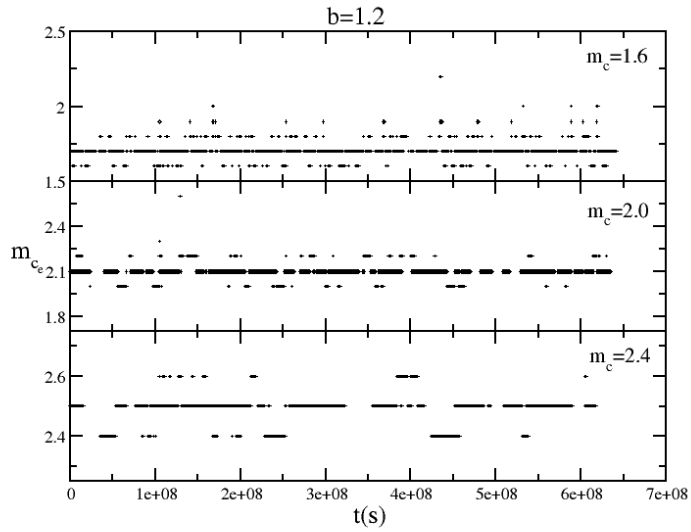


Figure 14 The time variation of the m_{c_e} value for the catalogue with b=1.2, thinned at three different m_c values.

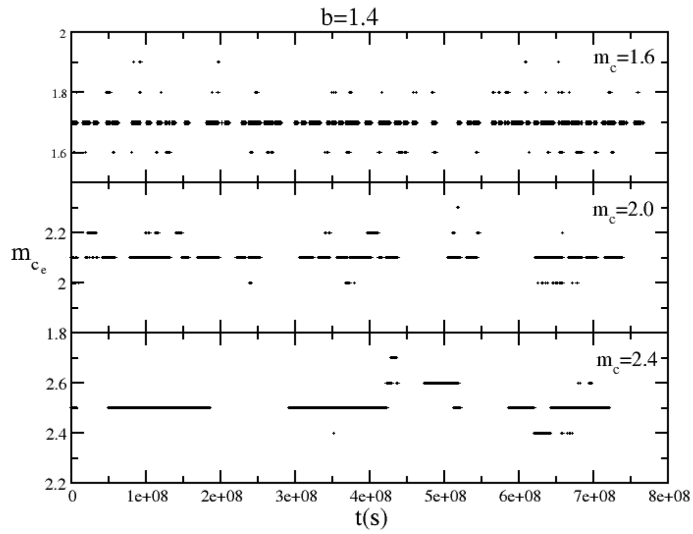


Figure 15 The time variation of the m_{c_e} value for the catalogue with b=1.4, thinned at three different m_c values.

that do not fall within the constraints of the evaluation of b_e, in particular, the required number of events N is greater than the number of events recorded in the catalogue.

Figs. 11 - 15 show the estimation of m_c for the same thinned catalogues. m_c estimates appear to be stable but are overestimated by a very small quantity, for each value of simulated b and completeness magnitude m_c. The error bars for b_e and m_{c_e} are not shown in order to make the graphs clearer. Furthermore, as this is a test on synthetic catalogues, there is no intention to evaluate the changes in b for forecasting purposes, but only to assess the used estimation method.

Fig. 16 shows the distribution of the b_e values for the different catalogues generated with the different b, confirming the tendency to underestimate b independently of the m_c value.

For comparison we evaluate the b values using the maximum curvature method for estimating m_c. Fig. 17 reveals that b_e significantly differs from the 'true' value

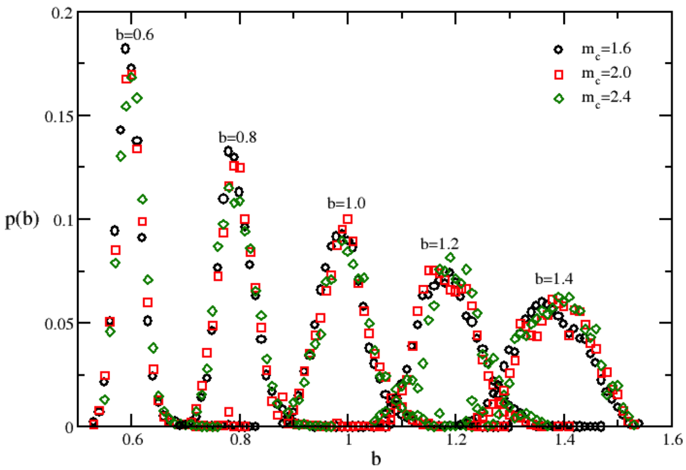


Figure 16 The distributions of the b_e value for the different catalogues thinned at the different m_c values.

in the case of b=1.0 and worsen as m_c increases. Indeed, the distribution's peak is large for m_c=1.6, bimodal for m_c=2.0, and sharply peaked at b_e=0.76 for m_c=2.4. Very similar behaviour is observed, but not reported here, for the other values of b.

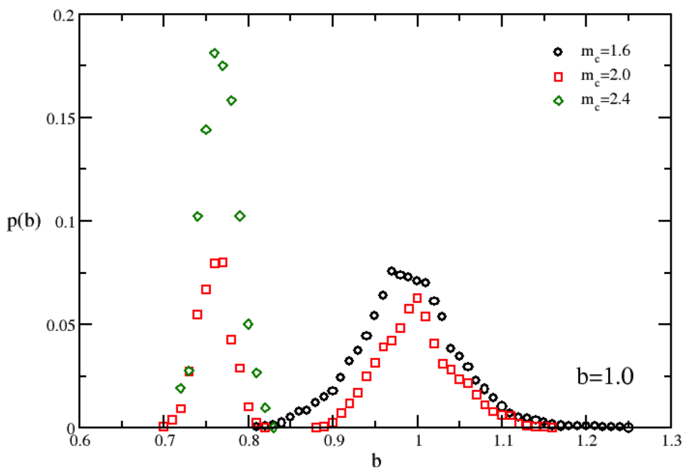


Figure 17 The distributions of the b_e in the case of b=1.0 and for the different m_c values, when m_c is estimated with the maximum curvature method.

3 Verifying the real-time discrimination of earthquake foreshocks and aftershocks

Analyzing the Amatrice-Norcia and Kumamoto sequences, Gulia and Wiemer (2019) demonstrated that variations in b can act as a discriminant between foreshocks and aftershocks. In practice, for the Italian Amatrice-Norcia sequence, they measured a reference value, b_r, for the background, considering the 4 years preceding the target earthquake $\mathcal{E}$. Then, removing events during the first 3 days after $\mathcal{E}$ because of short term aftershock incompleteness, they computed $\Delta b = b_{\mathcal{E}} - b_r$, where $b_{\mathcal{E}}$ is the b value computed using earthquakes ocurring after the event $\mathcal{E}$. The completeness magnitude m_c was estimated through the maximum-curvature method and the b value using a maximum-likelihood estimation on a sample of $N = 250$ events (consequently, the time series $b(t)$ was obtained with an element-wise moving window). Finally, b_r was extracted by calculating the median of all $b(t)$. The calculation of $b_{\mathcal{E}}$ follows the same rules but, because aftershock sequences are data-rich, the study used $N = 400$. However, Gulia and Wiemer (2019) also tested the results for slightly different values of N, and verified the robustness of their results.

3.1 Applying the c_v method - the comparison

Some authors questioned the results of Gulia and Wiemer (2019), arguing that fluctuations in the b value may not be a dependable indicator of stress in these instances. Instead, they could be attributed to a mixture of inconsistencies in the data and inefficiencies in estimation methods (Lombardi, 2022). Moreover, van der Elst (2021) confirms the findings of Gulia and Wiemer (2019), although at a reduced level. For this reason, to try to reduce estimation bias, we apply the Godano et al. (2023) method to the catalogue selected by Gulia and Wiemer (2019). More precisely, we evaluate m_c (using the Godano et al. (2023) method with a $c_{vt} = 0.93$) and b in the same windows of $N = 250$ events, sliding by one event at time and discarding all the windows with a number of events with $m \geq m_c$, larger than 50 and fulfilling the condition $m_{max} - m_c > 2$. Even if these conditions lead to discarding a large number of windows, the result of Gulia and Wiemer (2019) is confirmed. Indeed, both the average $\langle b \rangle$ value and its median b_M, before the occurrence of the Amatrice earthquake, are larger than the same values calculated after its occurrence and before the occurrence of the Norcia earthquake (Fig. 18). Interestingly, the difference between $\langle b \rangle$ and b_M, before the occurrence of the Amatrice earthquake, is large, revealing a significant skewness of the b distribution. Conversely, after the occurrence of the Amatrice earthquake, $\langle b \rangle - b_M$ assumes a very small value, indicating a Gaussian-like b value distribution.

Fig. 19 shows an example of two Gutenberg-Richter (GR) distributions for two different time windows (randomly chosen) before and after the Amatrice earthquake occurrence.

3.2 A blind algorithm

We now test an automatic algorithm for the real-time discrimination of earthquake foreshocks and aftershocks on the Italian catalogue available at the web-site http://terremoti.ingv.it/. We adopted the following algorithm:

1. Identify all the earthquakes with magnitude $m \geq 5.5$ (let us call them mainshocks);

2. For each one of them, we evaluate the aftershock radius following the Utsu and Seki (1954) formula $r = 0.05 \times 10^{0.5m} km$;

3. Identify all the events with a distance d smaller than r in the 4 years preceding the occurrence of the mainshock;

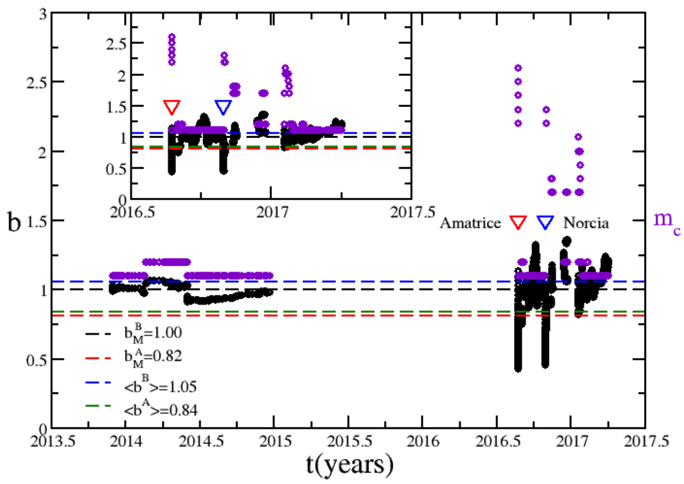


Figure 18 The b values (black circles) versus time. The dashed lines represent $\langle b \rangle$ and b_M (see legends for details) computed before (B) and after (A) the occurrence of Amatrice earthquake. The violet circles represent the m_c values. The triangles indicate the occurrence of the Amatrice and Norcia earthquakes. (Inset) Zoom of the main panel performed between 2016.5 and 2017.5.

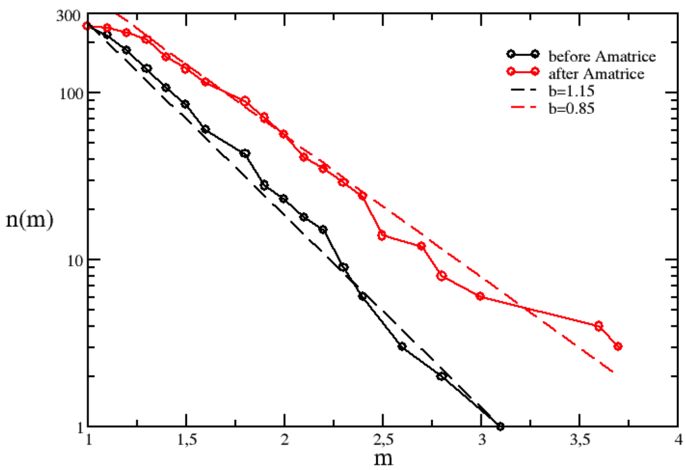


Figure 19 The Gutenberg-Richter distributions 4 years before and 1 year after the Amatrice earthquake occurrence. The dashed lines represent the fitted GR laws.

4. If the number n_b of events preceding the mainshock is smaller than 500 we double the value of r;

5. Identify all the events with a distance d smaller than r in 1 year following the occurrence of the mainshock;

6. We remove short-term aftershock incompleteness by means of the Helmstetter et al. (2006) method using the parameters optimized for Italy in Petrillo and Lippiello (2020);

7. If the number n_a of events following the mainshock is smaller than 500 we discard the mainshock from the analysis;

8. We evaluate the b and m_c values as a function of time following the previously described method.

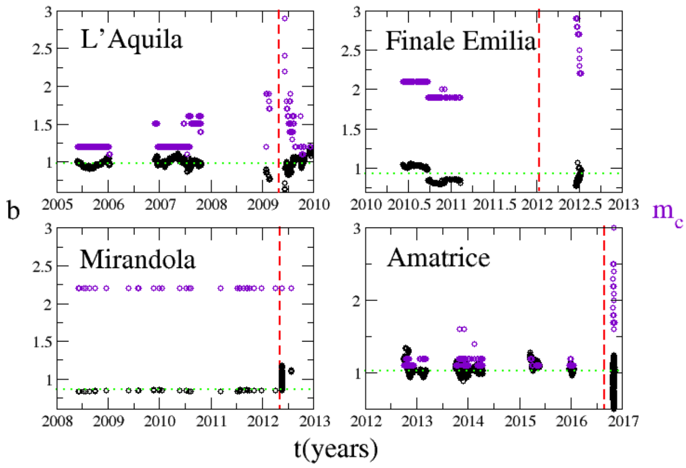


Figure 20 $b(t)$ and $m_c(t)$ for four of the mainshocks analysed here. Vertical red dashed lines represent the occurrence time of each mainshock. Horizontal green dotted lines represent the pre-event average b-value.

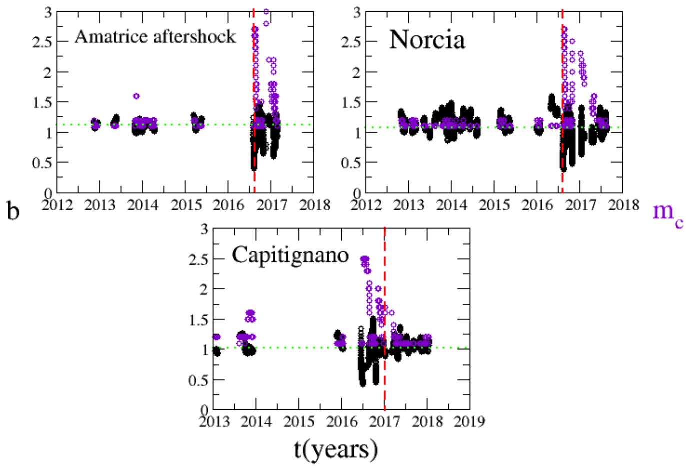


Figure 21 The $b(t)$ and $m_c(t)$ for the other three mainshocks analysed here. Vertical red dashed lines represent the occurrence time of each mainshock. Horizontal green dotted lines represent the pre-event average b-value.

On the basis of this algorithm, we identified mainshocks in the Italian catalogue. Item 3 has been applied only one time for the Finale Emilia earthquake. We discarded two mainshocks because n_a assumed values smaller than 500 for them. Both of these mainshocks ($m = 5.8$ and $m = 5.9$) occurred in the Aeolian Arc at a depth larger than 144 km. As a consequence, 7 earthquakes remain in the analysis: L'Aquila ($m = 6.1$), Finale Emilia ($m = 5.8$), Mirandola ($m = 5.6$), Amatrice ($m = 6.0$), an Amatrice aftershock ($m = 5.9$), Norcia ($m = 6.5$) and Capitignano ($m = 5.5$) close to L'Aquila. The information about the location, the magnitude, and the occurrence time of the mainshocks considered in the blind test is listed in Table 1.

An important deviation from Gulia and Wiemer (2019) is the use of the $M5.5 - 5.9$ earthquakes. This addition was decided in order to introduce more events in the blind test. A second deviation from Gulia and Wiemer (2019) is the use of the Utsu and Seki (1954) formula to identify the aftershocks and the background activity within a circular radius. This choice represents

Earthquake	Date	Location	Magnitude
L'Aquila	6 April 2009	42.42, 13.39	6.1
Finale Emilia	20 May 2012	44.80, 11.19	5.8
Mirandola	20 May 2012	44.85, 11.06	5.6
Amatrice	24 August 2016	42.70, 13.22	6.0
Amatrice aftershock	26 October 2016	42.90, 13.09	5.9
Norcia	30 October 2016	42.84, 13.11	6.5
Capitignano	18 January 2017	42.48, 13.28	5.5

Table 1 List of the mainshocks considered in the blind test including coordinates, occurrence time and magnitude.

the simplest one that can be performed in a blind algorithm, as it does not require the knowledge of the focal mechanism, the identification of the fault plane and its extension, and the localization of the mainshock on the fault. Moreover, the epicentral map (see supplementary information) of the chosen earthquakes reveals that, concerning the aftershocks, their selection corresponds to on-fault seismicity. Conversely, the selected background seismicity includes off-fault events. However, this does not represent a large difference with the Gulia and Wiemer (2019) method. Indeed, Gulia and Wiemer (2019) had to enlarge the investigated area in order to include a number of events sufficient to constrain the b_r value. The $b(t)$ and $m_c(t)$ values are shown in Figs. 20 and 21, and $\langle b \rangle$, b_M, and the b standard deviation σ_b before and after the occurrence of the 7 mainshocks are reported in Table 2.

Earthquake	before			after		
	$\langle b \rangle$	b_M	σ_b	$\langle b \rangle$	b_M	σ_b
L'Aquila	0.98	0.98	0.058	0.96	0.97	0.085
Finale Emilia	0.93	1.0	0.095	0.84	0.8	0.19
Mirandola	0.86	0.88	0.1	1.14	1.15	0.016
Amatrice aftershock	1.06	1.06	0.074	1.02	1.04	0.14
Amatrice	1.03	1.06	0.15	0.93	0.9	0.13
Norcia	1.07	1.06	0.16	1.0	1.02	0.092
Capitignano	1.02	1.03	0.16	1.07	1.07	0.087

Table 2 $\langle b \rangle$, b_M and σ_b for the 7 mainshocks analysed here, before and after their occurrence.

Earthquake	before			after		
	$\langle b \rangle$	b_M	σ_b	$\langle b \rangle$	b_M	σ_b
L'Aquila	0.99	0.99	0.064	0.95	0.96	0.082
Finale Emilia	0.93	1.0	0.094	0.86	0.85	0.062
Amatrice aftershock	1.11	1.10	0.057	0.94	0.94	0.14
Amatrice	1.02	1.04	0.16	1.02	1.05	0.16
Norcia	1.03	1.04	0.18	0.90	0.92	0.13
Capitignano	1.02	1.03	0.16	1.09	1.1	0.096

Table 3 $\langle b \rangle$, b_M and σ_b for the 6 mainshocks analysed considering restricted time period before (2 years) and after (0.5 years) the target mainshock.

Moreover, in order to test the robustness of the time parameters considered in the blind test, we show in Table 3 that halving the time period before and after the occurrence of the mainshock target does not change the results substantially. Note that halving the time period for the Mirandola earthquake is excluded from the analysis because the number of aftershocks is smaller than 500.

A t-test reveals that the b values before and after the occurrence of the mainshocks cannot be considered different at a 95% significance level. This implies that the discrimination between foreshocks and aftershocks cannot be performed using a blind algorithm.

Conclusions

We extensively tested the method proposed by Godano et al. (2023) for evaluating the completeness magnitude (m_c), which we referred to as the c_v method. The testing involved randomly generated catalogues with varying values of b and m_c, as well as simulated catalogues generated using ETAS models with fixed b and m_c values. In all cases, the method exhibited excellent performance. The distributions of the estimated values compared to the "true" values showed a supergaussian shape, centered around zero. Using the c_v method we then tested the results of Gulia and Wiemer (2019) for the Amatrice-Norcia earthquake sequence, confirming their results when using their catalogue, which represents a specific selection of the Italian catalogue. The use of a more reliable method in the estimation of m_c and b represents a stronger confirmation of their results, resolving doubts about possible biases introduced by an underestimation of m_c (Lombardi, 2022). However, when applying a blind algorithm to the Italian catalogue, no differences were found in the b values before and after the occurrence of the mainshock. This result indicates that: 1) the difference in b values is significant when an appropriate catalogue selection is made, supporting the notion that the b value serves as a reliable stress indicator. Specifically, if a genuine mainshock is imminent, the stress increases while the b value decreases; 2) the decrease in the b value cannot be detected using our blind approach and, consequently, cannot be utilized for real-time predictive purposes. Of course, other blind approaches could, conversely, confirm the Gulia and Wiemer (2019) results.

Acknowledgements

This study has been supported by the PRIN PNRR DIRECTIONS Project, No P20229KB4F, funded by the Italian Ministry of Education and Research. The authors thank two anonymous reviewers for their comments on this study.

Data and code availability

The catalogue used in this work can be downloaded at the web-site http://terremoti.ingv.it/.

Competing interests

The authors have no competing interests.

References

Aki, K. Maximum likelihood estimate of b in the formula $\log N = a - bM$ and its confidence limits. *Bull. Earthq. Res. Inst., Univ. Tokyo*, 43:237–239, 1965.

Amitrano, D. Brittle-ductile transition and associated seismicity: Experimental and numerical studies and relationship with the b value. *Journal of Geophysical Research: Solid Earth*, 108(B1): 2044, 2003. doi: 10.1029/2001JB000680.

Cao, A. and Gao, S. S. Temporal variation of seismic b-values beneath northeastern Japan island arc. *Geophysical Research Letters*, 29(9):48–1–48–3, 2002. doi: https://doi.org/10.1029/2001GL013775.

Console, R., Jackson, D., and Kagan, Y. Using the ETAS Model for Catalog Declustering and Seismic Background Assessment. *Pure and Applied Geophysics*, 167:819–830, 2007. doi: 10.1007/s00024-010-0065-5.

de Arcangelis L., Godano, C., R., G. J., and E., L. Statistical physics approach to earthquake occurrence and forecasting. *Physics Reports*, 628:1–91, 2016. doi: 10.1016/j.physrep.2016.03.002.

García-Hernández, R., D'Auria, L., Barrancos, J., Padilla, G. D., and Pérez, N. M. Multiscale Temporal and Spatial Estimation of the b-Value. *Seismological Research Letters*, XX:1–13, 2021. doi: https://doi.org/10.1785/0220200388.

Godano, C. A new method for the estimation of the completeness magnitude. *Physics of the Earth and Planetary Interiors*, 263:7 – 11, 2017. doi: https://doi.org/10.1016/j.pepi.2016.12.003.

Godano, C. and Petrillo, G. Estimating the Completeness Magnitude mc and the b-Values in a Snap. *Earth and Space Science*, 10(2), 2023. doi: 10.1029/2022EA002540.

Godano, C., Convertito, V., Pino, N. A., and Tramelli, A. An Automated Method for Mapping Independent Spatial b Values. *Earth and Space Science*, 9(6):e2021EA002205, 2022. doi: https://doi.org/10.1029/2021EA002205.

Godano, C., Petrillo, G., and Lippiello, E. Evaluating the incompleteness magnitude using an unbiased estimate of the b value. *Geophysical Journal International*, submitted, 2023.

Gulia, L. and Wiemer, S. The influence of tectonic regimes on the earthquake size distribution: A case study for Italy. *Geophysical Research Letters*, 37(10), 2010. doi: 10.1029/2010GL043066.

Gulia, L. and Wiemer, S. Real-time discrimination of earthquake foreshocks and aftershocks. *Nature*, 574:193–199, 2019. doi: 10.1038/s41586-019-1606-4.

Gutenberg, B. and Richter, C. Frequency of earthquakes in California. *Bulletin of the Seismological Society of America*, 34(4): 185–188, 1944.

Helmstetter, A. and Sornette, D. Foreshocks explained by cascades of triggered seismicity. *Journal of Geophysical Research: Solid Earth*, 108(B10):2457, 2003. doi: 10.1029/2003JB002409.

Helmstetter, A., Kagan, Y. Y., and Jackson, D. D. Comparison of Short-Term and Time-Independent Earthquake Forecast Models for Southern California. *Bulletin of the Seismological Society of America*, 96(1):90–106, 2006. doi: 10.1785/0120050067.

Kagan, Y. Y. Aftershock zone scaling. *Bulletin of the Seismological Society of America*, 92(2):641–655, 2002. doi: 10.1785/0120010172.

Kamer, Y. and Hiemer, S. Data-driven spatial b value estimation with applications to California seismicity: To b or not to b. *Journal of Geophysical Research: Solid Earth*, 120(7):5191–5214, 2015. doi: https://doi.org/10.1002/2014JB011510.

Lippiello, E., Giacco, F., Arcangelis, L. d., Marzocchi, W., and Godano, C. Parameter estimation in the ETAS model: Approximations and novel methods. *Bulletin of the Seismological Society of America*, 104(2):985–994, 2014. doi: 10.1785/0120130148.

Lombardi, A. M. Anomalies and transient variations of b-value in Italy during the major earthquake sequences: what truth is there to this? *Geophysical Journal International*, 232(3): 1545–1555, 10 2022. doi: 10.1093/gji/ggac403.

Lombardi, M. and Marzocchi, W. The ETAS model for daily forecasting of Italian seismicity in the CSEP experiment. *Annals of Geophysics*, 53, 2010. doi: 10.4401/ag-4848.

Mignan, A. and Woessner, J. Estimating the magnitude of completeness for earthquake catalogs. *Community Online Resource for Statistical Seismicity Analysis*, 2012. doi: 10.5078/corssa-00180805.

Ogata, Y. Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes. *J. Am. Stat. Assoc.*, 83(401): 9–27, 1988. doi: 10.2307/2288914.

Ogata, Y. Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics*, 50(2): 379–402, 1998. doi: 10.1023/A:1003403601725.

Ogata, Y. Seismicity Analysis through Point-process Modeling: A Review. *pure and applied geophysics*, 155(2-4):471–507, 1999. doi: 10.1007/s000240050275.

Ogata, Y. and Katsura, K. Analysis of temporal and spatial heterogeneity of magnitude frequency distribution inferred from earthquake catalogues. *Geophysical Journal International*, 113 (3):727–738, 1993. doi: 10.1111/j.1365-246X.1993.tb04663.x.

Ogata, Y. and Zhuang, J. Space–time ETAS models and an improved extension. *Tectonophysics*, 413(1–2):13 – 23, 2006. doi: http://dx.doi.org/10.1016/j.tecto.2005.10.016.

Ohmura, A. and Kawamura, H. Rate- and state-dependent friction law and statistical properties of earthquakes. *EPL (Europhysics Letters)*, 77(6):69001, 2007. doi: 10.1209/0295-5075/77/69001.

Papadopoulos, G., Minadakis, G., and Orfanogiannaki, K. *Short-Term Foreshocks and Earthquake Prediction*, chapter 8, pages 125–147. American Geophysical Union (AGU), 2018. doi: https://doi.org/10.1002/9781119156949.ch8.

Papadopoulos, G. A. Long-term accelerating foreshock activity may indicate the occurrence time of a strong shock in the Western Hellenic Arc. *Tectonophysics*, 152(3–4):179 – 192, 1988. doi: http://dx.doi.org/10.1016/0040-1951(88)90044-3.

Papadopoulos, G. A., Charalampakis, M., Fokaefs, A., and Minadakis, G. Strong foreshock signal preceding the L'Aquila (Italy) earthquake (M_w 6.3) of 6 April 2009. *Natural Hazards and Earth System Science*, 10(1):19–24, 2010. doi: 10.5194/nhess-10-19-2010.

Petrillo, G. and Lippiello, E. Testing of the foreshock hypothesis within an epidemic like description of seismicity. *Geophysical Journal International*, 225(2):1236–1257, 12 2020. doi: 10.1093/gji/ggaa611.

Petrillo, G. and Lippiello, E. Incorporating Foreshocks in an Epidemic-like Description of Seismic Occurrence in Italy. *Applied Sciences*, 13(8), 2023. doi: 10.3390/app13084891.

Petrillo, G. and Zhuang, J. The debate on the earthquake magnitude correlations: A meta-analysis. *Scientific reports*, 12(1): 20683, 2022. doi: 10.1038/s41598-022-25276-1.

Petrillo, G. and Zhuang, J. Verifying the magnitude dependence in earthquake occurrence. *Physical Review Letters*, 131(15): 154101, 2023. doi: 10.1103/PhysRevLett.131.154101.

Pino, N. A., Convertito, V., Godano, C., and Piromallo, C. Subduction age and stress state control on seismicity in the NW Pacific subducting plate. *Sci. Rep.*, 12, 2022. doi: 10.1038/s41598-022-16076-8.

Roberts, N. S., Bell, A. F., and Main, I. G. Are volcanic seismic b-values high, and if so when? *Journal of Volcanology and Geothermal Research*, 308:127–141, 2015. doi: 10.1016/j.jvolgeores.2015.10.021.

Rydelek, P. A. and Sacks, I. S. Testing the completeness of earthquake catalogs and the hypothesis of self-similarity. *Nature*, 337:251–253, 1989. doi: 10.1038/337251a0.

Scholz, C. The frequency-magnitude relation of microfracturing in rock and its relation to earthquakes. *Bull. seism. Soc. Am.*, 58: 399–415, 1968. doi: 10.1785/BSSA0580010399.

Taroni, M., Zhuang, J., and Marzocchi, W. High-Definition Mapping of the Gutenberg–Richter b-Value and Its Relevance: A Case Study in Italy. *Seismological Research Letters*, XX:1–7, 2021. doi: https://doi.org/10.1785/0220210017.

Tormann, T., Wiemer, S., and Mignan, A. Systematic survey of high-resolution b value imaging along Californian faults: Inference on asperities. *Journal of Geophysical Research: Solid Earth*, 119 (3):2029–2054, 2014. doi: 10.1002/2013JB010867.

Tramelli, A., Troise, C., De Natale, G., and Orazi, M. A new method for optimization and testing of microseismic networks: an application to Campi Flegrei (Southern Italy). *Bulletin of the Seismological Society of America*, 103(3):1679–1691, 2013. doi: 10.1785/0120120211.

Tramelli, A., Godano, C., Ricciolino, P., Giudicepietro, F., Caliro, S., Orazi, M., De Martino, P., and Chiodini, G. Statistics of seismicity to investigate the Campi Flegrei caldera unrest. *Scientific reports*, 11(1):7211, 2021. doi: 10.1038/s41598-021-86506-6.

Utsu, T. and Seki, A. A relation between the area of aftershock region and the energy of mainshock. *J. Seismol. Soc. Jpn.*, 7:233 – 240, 1954.

van der Elst, N. J. B-Positive: A Robust Estimator of Aftershock Magnitude Distribution in Transiently Incomplete Catalogs. *Journal of Geophysical Research: Solid Earth*, 126(2):e2020JB021027, 2021. doi: 10.1029/2020JB021027.

Wiemer, S. A Software Package to Analyze Seismicity: ZMAP. *Seismological Research Letters*, 72(3):373–382, 05 2001. doi: 10.1785/gssrl.72.3.373.

Wiemer, S. and Wyss, M. Mapping the frequency-magnitude distribution in asperities: An improved technique to calculate recurrence times? *J. Geophys. Res.*, 102:15,115–15,128, 1997. doi: 10.1029/97JB00726.

Wiemer, S. and Wyss, M. Minimum Magnitude of Completeness in Earthquake Catalogs: Examples from Alaska, the Western United States, and Japan. *Bulletin of the Seismological Society of America*, 90(4):859–869, 2000. doi: 10.1785/0119990114.

Wiemer, S. and Wyss, M. Mapping spatial variability of the frequency-magnitude distribution of earthquakes. *Adv. Geophys.*, 45:259–302, 2002. doi: 10.1016/S0065-2687(02)80007-3.

Wyss, M. Towards a Physical Understanding of the Earthquake Frequency Distribution. *Geophysical Journal of the Royal Astronomical Society*, 31(4):341–359, 1973. doi: 10.1111/j.1365-246X.1973.tb06506.x.

Y., O. Estimation of the Parameters in the Modified Omori Formula for Aftershock Frequencies by the Maximum Likelihood Procedure. *Journal of Physics of the Earth*, 31(2):115–124, 1983.

Zhuang, J. Next-day earthquake forecasts for the Japan region generated by the ETAS model. *Earth Planets and Space*, 63: 207–216, 2011. doi: doi.org/10.5047/eps.2010.12.010.

Zhuang, J. Long-term earthquake forecasts based on the epidemic-type aftershock sequence (ETAS) model for short-term clustering. *Research in Geophysics*, 2, 2012. doi: 10.4081/rg.2012.e8.

Zhuang, J. and Touati, S. Stochastic simulation of earthquake catalogs. *Community Online Resource for Statistical Seismicity Analysis*, 2015.

Zhuang, J., Ogata, Y., and Vere-Jones, D. Analyzing earthquake clustering features by using stochastic reconstruction. *Journal of Geophysical Research: Solid Earth*, 109(B5), 2004. doi: 10.1029/2003JB002879.

Rebuttal letter

We thank both reviewers for their work that have helped us to improve the manuscript. We have done our best to answer their questions.
Some confusion came from the fact that we sometimes wrongly used the work "anomalies" instead of "heterogeneities" which definitely didn't help the understanding of our point. We refer to "anomalies" for DAS signals that are not consistent with the wave propagation (typically too small spatially for a given frequency). Heterogeneities are regions with mechanical properties different from a background.
Another significant source of confusion arose from the misconception that the numerical tests were designed to replicate the data presented in Section 3. However, this is not the case. The paper's methodology involved initially developing the numerical model, followed by seeking an actual data example to illustrate our findings. We have made efforts to clarify this aspect to eliminate any potential confusion.

1 Reviewer 1

- 1 17-18 might be simplified
  we have tried to improve this part
- 2 31-32 hard to read
  we have tried to improve this part.
- 3 37 not a density
  linear density?
- 4 44 not clear which ones !
  we mean : 1: simply pulled in underground conduits 2: the FO is decoupled from the outer jacket
- 5 45 not clear
  We have tried to improve that
- 6 61-62 here there is something unclear. It has never be said before that DAS is a measurement over an extended distance, then the opposition of strain and point measurement is not straightforward
  Our mistake: a few words were missing which makes the sentence have no meaning. We added "... particle displacement or velocity point measurements." Sorry about that.
- 7 74 not sure of that, do you have a reference ? Their are also Brioullin and Raman non elastic scattering. but Rayleigh scattering is what is analyzed

  We removed the word "dominant" which is not critical for the discussion
- 8 79 strain (Dl/l), where elongation is Dl
  done
- 9 83 difficult to read
  changed
- 10 86 exponent -12 lacking Hz should not be italic I think the distance is also required
  Done
- 11 95 receiver spacing and gauge length are independant I think it was fixed in first version of Febus but it is not always the case

  Corrected. The gauge length here was actually 9.6 m

- 12 102 add ref for this formula. also is xi+1 ¿xi the dt should be negative: the optical wave reflect on xi+1 and interfere at xi with a delay, then the optical delay is on ux and dt should be explicitely given in function of the distance difference and c. this formula need to be explained with more detailled as it is the basis of the paper

  We changed the formula and added a reference
- 13 106 the rayleigh scattering is not evaluated in DAS
  done
- 14 113 epicenter position?
  Done
- 15 124 mixed between singular and plural please correct
  corrected
- 16 126 is that the black lines ? it is perturbating to refer to 900 sensors better add between the distances 8000? to 10000 m because no reference of the sensor number is given in fig1
  Indeed, it is the black lines. We have changed the text to refer to the black lines.
- 17 127 very difficult to understand color of the filter seems to be inverted (red signals contains higher frequencies (even if seen in space)) Also the waveforms seems to be normalized in the representation, then you can refer to attenuation.
  We admit it is not obvious, but the color scale is not inverted. The main reason for this confusing impression is the anomalies that are dominating the signal. We have added a magenta box in fig1(c) to highlight one example of signal we wish to show.
- 18 129-131 not clear
  We have tried to improve this part
- 19 144 Is that related to the method you use to solve the elasto dynamic problem? I do not think it is written where it should be.
  Maybe we are missing the point of this comment. When defining the problem, we need to specify the boundary conditions used. Here, we neither use Dirichlet nor Neumann boundary conditions, but absorbing boundaries. Those are meant to absorb any outgoing waves. We are not stating yet what method we are using to implement those boundary conditions.
- 20 151 does look like rather a monopole /explosion than a usual vector point source
  This is a standard expression for a moment tensor point source. See, for example, Dahlen & Tromp (1998) Eq. 5.85 page 166. Please note that the gradient operator applies to the Dirac distribution.
- 21 155 indeed see comment at line 144-145
  See above answer
- 22 156 constante?
  yes
- 23 160 please check compatibility with eq 2
  done
- 24 160 mainly dt disappeared in first term
  It did because we are using v instead of u. But, this is a good point, we change Eq. 2 to be the same as Eq 6.
- 25 161 not in the equation 6
  Not sure about the point of this comment. Δ_x is in Eq. 6.

- 26 172 vp is too low, see brocher relation to fix that
  Indeed, we used a V_P that we can find in dry sand or topsoil (The first author had in mind a FO in volcano hashes at the very beginning, and it unfortunately stayed that way). Because changing that would imply re-running everything, re-do all the figures without changing the visual aspect of the figures nor any conclusions of this work, we didn't change that value. We removed the sentence "representative of subsurface layers" as it is not accurate.

- 27 175 I m lost here. You never mentionned 2 fibers before. Also fibers are on the surface not in the bulk as in your simulation. Then shoul I understand that you simulated only surface waves or disregard the effect of surface wave reflection on your analysis?
  Indeed, the numerical experiment presented here should not be seen as an attempt to model the sec. 3 real data. Nevertheless, such a simple modeling is sufficient to reach our point: the effect of small scale heterogeneities on DAS measurement. And yes, we are using two FO cables in our numerical experiment to make our point.

- 28 178 then it is not to mimic the FO cables?
  See the above answer: it is to illustrate FO cables in general, not specifically the one of the previous observations in Sec. 3.

- 29 fig 3 can we see the effect on S wave too ?
  Because the chosen source is an explosion, there is no ballistic S-wave. There are S waves in the scattered wavefield after the ballistic P-wave, but they are difficult to isolate.

- 30 188 Why rock are discretized by only 2 cells? Don t you require at least 6 cells for an heterogeneity to ensure that the errors is not numerical? The diffracted field by rock is it well sampled close the rock? Perhaps in farefield but certainly not in nearfield. I think the convergence close to the diffractors should be shown in supplementary at least. You also need to discuss the error in displacement and in strain, I think they differ from 1 order due to the derivative. What is the ratio impose between element length and lambda_min?
  A general fact of the dynamic is that, to obtain good quality result, it is enough to mesh the geometry (discontinuities) and the minimum wavelength. This is not true for the static case, for which singularities of the solution near complex geometry implies a finer and finer mesh. Knowing that, a single element per stone would be enough. We have 4 elements/stone here because the meshing tool (GMSH) we use here is not smart enough to join the 4 elements into 1. Another meshing tool (CUBIT and its successor, commercial) would have been able to generate a mesh with a single element par stone. The polynomial degree used here is 6, which means each element has 7x7 "points" (Gauss-Lebatto-Legendre collocated points). Such a degree is enough to accurately model any wavefield using 1 element per minimum wavelength. Here, because we have to mesh the stones, the elements are very small compared to the minimum wavelength (in a smooth media, the elements could be 15 times larger, the wavefield would still be accurate). The wavefield is therefore seriously over-sampled, which, together with the well-known accuracy of the SEM, ensures we have no accuracy issue here. If you are interested in knowing more about the spectral element convergence, there are many studies, but one of the first is Seriani & Priolo (1994).

- 31 190-194 Why not introducing g dot in 5 and use v and strain rate? The explanation even if right is complicating the paper.
  We have thought about that and what you suggest was indeed a considered option. But this discussion about g depends on the nature of the source (g is a step function for an earthquake, a Dirac for a triggered explosion etc). Because in the end it doesn't change much the results and because its application is not tied to the source nature, we decided to keep it that way.

- 32 197 one case 2 variables this is confusing because there are 2 fibres,
  We changed that

- 33 197 we only see from 26 s, change the description of first
  The signal is zero before 26s, this is the first arrival. It took about 25s for the P waves to reach the FO1 (+2s to account for the g zero time).

- 34 202 this is only from here I can understand what anomalies are in section 3 However I am not very convinced. In fig3 du, the balistic wave has a dicontinuous amplitude over the different chanel and it iseems more related to channel normalization. The amplification might be related to local site effect, and here you are in a bulk.
  We are not sure what is meant here. About the normalization: it is done by maps, not by traces, so that the relative amplitudes within the maps are preserved. Moreover, the cross-section for a given time step in Fig 5 (formerly fig 4) right panel shows clear "glitches" (anomalies) and only for the DAS "strain" measurement, not for the displacement (fig 5 left panel). This observation can indeed be called a "site effect", but then it is an effect only for a strain measurement, not for a displacement measurement at the same location. Classical "site effects" are ringing effects that still obey a dispersion relation (they appear only in some frequency bands), which is not the case here.

- 35 Fig 6 a portion of
  The caption has been updated

- 36 Fig 7 It is not clear that the effect is lower on FO2 as everyting is normalized. Can you show it also not normalized?
  The traces are now normalized by a common factor (the maximum amplitude of FO1) which makes the amplitude comparable.

- 37 267 I am not sure about the shortcut.
  "Similarly" has been removed

- 38 271 better use index at left ?
  This is possible. Nevertheless, this notation used here is very common for many two-scale homogenization works.

- 39 310 Eq?
  Done

- 40 Fig 8 Is it a 1D material ? 1D homogeneisation?
  Fig 9 (previously Fig 8) is a cross-section of the 2-D homogenized solution. The method employed is the one described in Section 5. See the beginning of section 5.3.

- 41 380 The effect of filtering is not given explicitly in section 3. Is that for changing wavelength?
  Exactly. Following the dispersion relation, changing the frequency should change the wavelength. If it is not the case, the observed signal is an anomaly. We have tried to improve section 3 accordingly.

- 42 381 Not clear, DAs measurement shows anomalies. There are not affected by. Or do you mean the diffractors? Did you verify where the vertical lines stand that there are shallow heterogeneities?
  This is our mistake. We meant "display anomalies". This has been corrected and the paragraph rewritten.

- 43 402 I disagree, see site effects due to free surface topography.
  We are not sure why it is a disagreement. Indeed, free surface fine-scale topography has a similar effect compared to fine-scale heterogeneities regarding displacement versus strain. See Capdeville & Marigo (2013).

2 Reviewer 2

- 1. Page 2, Line 39 : What do the authors mean by cable coating or coupling? Kindly elaborate.
  We replaced "coating" by "cable jacket" and rephrased the sentence as: "The attachment of the Fiber Optic (FO) cable to both its cable jacket and the ground is known to strongly affect DAS measurements "

- 2. Page 3, Line 80 : What does the parameter ξ signify in Equation 1? Explained after the equation: " ξ is assumed constant, typically 0.79 in single-mode fiber, and acknowledges that straining the fiber also implies a proportional change in its refractive index". The reader can refer to the reference provided for the equation to get more details.

- 3. Page 3, Line 88 : Kindly elaborate on what the authors mean to convey in the statement "But some experiments on short fiber segments exploited data acquired at several kHz". This sentence is indeed not critical and was removed.

- 4. Page 4, Line 111 : Please provide further details (Date, Time, Coordinates) of the 2.7Mw aftershock recorded by the DAS fiber optic cable Done

- 5. Page 5 : Please specify the lengths of the two arms of the V-shaped network in Figure 1. The data shown in Figure 1 correspond to which time of the year approximately?

  Done. Figure 1 corresponds to the date the aftershock studied i.e. Nov. 23rd, 2019

- 6. Page 4, Line 123 : Is there a specific reason for choosing the 4 high-frequency cutoff? Why was it not just 5, 10, 15, 20 and 25?

  We tried different values and found that the chosen high-frequency cutoff frequencies to better illustrate the attenuation of the seismic waves with respect to the anomalies.

- 7. Page 4, Line 130-131 : The authors state that they do not observe the effect of filtering over one or few channels. Is it possible for the authors to mark at least one such region in Figure 1c to back this statement?
  We added a white box in Fig 1(b) and a cyan box in Fig 1(c) to display one example.

- 8. Page 4, Line 132: What do the authors mean by 'environment of the fiber'. Kindly elaborate.
  We mean anything near the fiber. We elaborated on the sentence

- 9. Page 4, Line 133 : When do the two anomalies, phase unwrapping errors and laser frequency drifts, occur in DAS acquisition?

  We provided examples: "Two well-known sources of anomalies inherent to DAS technology are phase unwrapping errors (e.g. something directly touching the OF and causing a sudden large strain) and laser frequency drifts (e.g. if the laser of the DAS is affected by external vibrations)"

- 10. Page 6, Line 139 : What do the authors mean by mechanical property anomalies? Please elaborate on this term before the paper delves into the numerical analysis.
  Sorry, we meant "heterogeneities" instead of "anomalies". This error is embarrassing as it is very confusing. Indeed, we observe anomalies in DAS measurements that we propose to link to small-scale heterogeneities in our work.

- 11. Page 6, Line 159 : What is pulse width effect? Kindly elaborate.
  The pulse width effect is explained just before section 3 and is expected to be small. We have moved the sentence "We ignore the pulse width effect" there for clarity.

- 12. Page 6, Line 160: In Equation 6, does ‘z’ represent depth of the medium? If not, the reviewer requests the authors to replace ‘z’ with some other variable.
  The ’z’ component represents more a horizontal component than a depth. Nevertheless, the 2-D modeling we are using here only represents partly a 3-D realistic setting and the $(\hat{x}, \hat{z})$ are just the basis vector of this 2-D setting, which is arbitrary and a common practice for numerical modeling. Moreover, we have used z instead of y to avoid confusion with the homogenization $\mathbf{y}$ variable. We have added a sentence to make sure there is no confusion.

- 13. Page 7 : Figure 2 is extremely confusing to the readers. As the y-axis in ‘Figure 2a’ is marked ‘z’, it is presumed that it represents the depth of the region of simulation. If so, then the origin of the region should be at top left corner instead of bottom left corner and the depth should be represented as -10km, -20km and so on. In that case it is easy for the reader to perceive that the explosion occurs at a depth of 30km from the top surface and the measurements are being recorded at a distance of 10km approximately. However, DAS fibre at a depth of 10km does not make sense. If ‘z’ is not the depth of the medium, then how is the simulation carried out? Is the source also assumed to be present at the same level as the DAS? The authors need to clarify on these points.
  We have added a text at the beginning of section 2 to clarify this point. As mentioned earlier, we present only a 2-D model in an infinite plate that can not represent many aspects of the 3-D real setting. We have added a section in the discussion too, about the limitations of the presented modeling.

- 14. Page 7, Line 178 : Either FO1 is directly in contact with 6 stones and FO2 is with none (according to current Figure 2b) or Figure 2b needs to be checked if there are 6 or 7 stones present on FO1.
  Indeed, Fig. 2b represents only a portion of the whole area. This portion is presented with a dotted line square in Fig. 2a. In Fig. 2b, only 6 stones are visible, but a 7th one is touching FO1. We added a sentence in the text to be more precise.

- 15. Page 7, Line 179 : When the authors write, 20km away from the cables, do they mean at a horizontal distance of 20km or a depth of 20km.
  As explained above, it is a 2-D modeling. We mean 20km away in the Ω plane. If we think of this 2-D modeling as a top view of a 3-D modeling, it is indeed 20km in the horizontal distance.

- 16. Page 8, Line 180-181 : Please provide a figurative representation of the Ricker source-time function showing the central and maximum frequency used to simulate the explosive source.
  A figure has been added. (Ricker functions are quite commonly used for numerical modeling in seismology)

- 17. Page 8, Line 186 : Kindly elaborate on what is GMSH. Also, please provide it’s full-form.
  GMSH is an open-source software. We added a sentence saying so. The necessary information to know more about it comes with the provided reference.

- 18. Page 8 : Is it possible for the authors to add a subplot of FO1 showing the location of the 7 stones right below both the subplots of Figure 3? It will improve the understanding of the reader to compare the not-so strong and strong glitches on the displacement and strain fields respectively with the actual location of heterogeneities on FO1.
  Done.

- 19. Page 8, Line 201-202 : Please mark the regions in Figure 1c, where authors think that they observe this phenomenon.
  done with a white box in Fig. 1(c)

- 20. Page 8, Line 203 : Kindly elaborate on trace collection representation.
  Added in the figure caption

- 21. Page 9 : It will also be good if the authors can mention the exact distance of these 6/7 stones on the x-axis with respect to the origin. It will help to match these distances with the peaks observed in the blackline in Figure 5.
  The positions of the stones along the FO as been marked by vertical gray lines in Fig 6 (formerly 5) and 8 (formerly 7)/

- 22. The authors are requested to show Figure 2b in the range of 6km to 14km so that this representation is consistent with their observations shown in Figure 3. It is also advisable to mark a small region around FO1 and FO2 which signifies what they show in Figure 4 and Figure 6.
  It is not easy the change the x-range without making the figure difficult to visualize with a standard resolution. Instead, we have added two panels showing the stone locations just below the traces collections.

- 23. The captions for Figure 4 and 6 are incomplete.
  Indeed, thank you. We have completed the caption

- 24. Page 10: It will be interesting to mark the exact location of stones present in FO1 in Figure 7. Also, it is visible in Figure 1b that there are 2 stones present extremely near to FO2. One is at z = 30km (below FO2) and x is between 9 to 10km and the other is on FO1. It will be interesting if the authors can try to observe and distinguish between these two stones from the strain amplitudes shown in Figure 7.
  We have added vertical lines in Fig 8 (formerly 7) to display the stone locations.

- 25. Page 12 : I request the authors to use numbers instead of bullets to enlist the main points of the development of homogenization theory.
  Done

- 26. I request the authors to add the word 'equation' before referring to any equation in the text to provide better readability to the manuscript.
  Done

- 27. Page 14, Line 327-330: The authors mention that a spatially smaller or larger heterogeneity has a similar amplitude effect on the strain. As per the understanding of the reviewer, the authors, in the present numerical simulation have considered the stones of the same size to represent heterogeneties around FO1 and FO2. It will be extremely interesting to show another set of simulations where the stones of different sizes are cautiously placed near FO1 and FO2 and plot figures similar to the left panels of Figures 3 and 4 and Figure 5. This set of simulations and corresponding figures will provide a better understanding of the statements mentioned in these lines.
  Indeed, it is a non-intuitive behavior. We have added a reference in which this effect is illustrated.

- 28. Page 15, Line 348-350: Kindly briefly describe the method to obtain effective elastic tensor so that the readers can refer to Browaeys and Chevrot (2004) for its details. The added explanation will provide a better understanding to Figure 10 for the readers.
  Browaeys and Chevrot are used only to compute the nearest isotropic elastic tensor $\mathbf{c}^{\text{iso}}$ to an arbitrary tensor $\mathbf{c}$. It is a simple method (and not a homogenization method) and we use it to represent the effective tensor obtained with the homogenization method Capdeville et al. (2010) described in the paper.

- 29. Page 15: Figure 8 caption- Change upper right panel to upper left panel. Also, the upper panels of Figure 8 can be combined to show the velocity in the original media and effective media in the same subplot as the authors have represented density.
  We made two separated plots for V_S and $V_S^{*\text{iso}}$ because the difference of amplitude makes the $V_S^{*\text{iso}}$ difficult to visualize if plotted in a single graph.

- 30. Page 16 : How many iterations are carried out by the authors to get the close match between the simulations in the original media and effective media shown in Figure 10b and 10d.

  It is not clear what is meant by iterations here as it is not an iterative method. There are only 3 steps to obtain the effective traces shown in Fig 11 (formerly 10):

  1. obtain the effective media and the correctors with the homogenization tool presented in section 5.1
  2. run SEM in the effective media (it is a standard SEM run) to obtain Fig 11 a and b
  3. apply the correctors to obtain Fig 11 c and d

- 31. Page 17, Line 388: Please provide appropriate references that show the numerical evaluation of small-scale heterogeneties on DAS measurements.

  This is the point of the paper. To our knowledge, there is no other.

- 32. It will be interesting if the authors try to simulate the 2.7Mw earthquake mentioned in Figure 1 assuming as an effective media and comparing it with the records that they have from the DAS measurements in the original media. This will further support the fact that by circumventing the effort that researchers may have to go through to characterize the heterogeneities in the substructure of the Earth to obtain realistic estimates of strain and rotations, the homogenization method can provide reliable results.

  This is a difficult request because we do not have a good model (at least good enough) of the area. This would nevertheless be possible with limited accuracy and gathering models and geology of the area and then using a 3-D SEM tool. The amount of work required to achieve such a goal makes us acknowledge that we can not comply for now, but it will be for a future work.

  Nevertheless, the requests show a problem in the writing of our paper. Indeed, it indicates that the reviewer thinks that we are trying to model the data shown in section 3, which we are not. The data are just here to illustrate the effect we are proposing to explain.

References

Capdeville, Y. & Marigo, J.-J., 2013. A non-periodic two scale asymptotic method to take account of rough topographies for 2-d elastic wave propagation. *Geophys. J. Int. 192*(1), 163–189.

Dahlen, F. A. & Tromp, J., 1998. *Theoretical Global Seismology.* Princeton University Press. NJ.

Seriani, G. & Priolo, E., 1994. Spectral element method for acoustic wave simulation in heterogeneous media. *Finite elements in analysis and design 16*(3-4), 337–348.

doi:10.26443/seismica.v3i1.980

SEISMICA

Detection of slow slip events along the southern Peru - northern Chile subduction zone

J. Jara * [1], **R. Jolivet** [1,2], **A. Socquet** [3], **D. Comte** [4,5], **E. Norabuena** [6]

[1]Laboratoire de Géologie, Départament de Géosciences, École Normale Supérieure, PSL Research University, CNRS UMR 8538, Paris, France, [2]Institut Universitaire de France, Paris, France, [3]Université Grenoble Alpes, Université Savoie Mont Blanc, CNRS, IRD, IFSTTAR, ISTerre, Grenoble, France, [4]Departamento de Geofísica, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Blanco Encalada 2002, Santiago, Chile, [5]Advanced Mining Technology Center, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Av. Tupper 2007, Santiago, Chile, [6]Instituto Geofísico del Perú, Lima, Perú

Author contributions: *Conceptualization*: A. S., J. J, R. J.. *Methodology*: J. J., R. J. *Formal Analysis*: J. J. *Investigation*: J. J. *Resources*: J. J., A. S., R. J., D. C., E. N. *Writing - original draft*: J.J. *Writing - Review & Editing*: J. J., R. J., A. S., D. C.

Production Editor:
Gareth Funning
Handling Editor:
Wenbin Xu
Copy & Layout Editor:
Ethan Williams

Signed reviewer(s):
Piero Poli, Jeffrey Freymueller

Received:
11 May 2023
Accepted:
3 May 2024
Published:
7 June 2024

Abstract Detections of slow slip events (SSEs) are now common along most plate boundary fault systems globaly. However, no such event has been described in the south Peru - north Chile subduction zone so far, except for the early preparatory phase of the 2014 Iquique earthquake. We use geodetic template matching on GNSS-derived time series of surface motion in Southern Peru - Northern Chile to extract SSEs hidden within geodetic noise. We detect 24 events with durations ranging from 17 to 36 days and magnitudes from M_w 5.4 to 6.2. Our events, analyzed from a moment-duration scaling perspective, reveal values consistent with observations reported in other subduction zones. We compare the distribution of SSEs with the distribution of coupling along the megathrust derived using Bayesian inference on GNSS- and InSAR-derived interseismic velocities. From this comparison, we obtain that most SSEs occur in regions of intermediate coupling where the megathrust transitions from locked to creeping or where geometrical complexities of the interplate region have been proposed. We finally discuss the potential role of fluids as a triggering mechanism for SSEs in the area.

Resumen Hoy en día, las detecciones de eventos lentos (SSEs, por sus siglas en inglés) son comunes a lo largo de la mayoría de los sistemas de fallas activas a una escala global. Sin embargo, hasta ahora, no se han reportado eventos de este tipo en la zona de subducción del sur del Perú y norte de Chile (10^oS-24^oS), exceptuando aquellos ocurridos durante la fase de preparación del terremoto de Iquique de 2014. En el presente trabajo, nosotros utilizamos una técnica conocida como "Template Matching" en series temporales de desplazamiento medido por datos GNSS (Global Navigation Satellite System, GNSS por sus siglas en inglés) en el sur del Perú y el norte de Chile, para extraer la firma de eventos lentos asísmicos ocultos en el ruido geodésico. Nosotros detectamos 24 eventos asísmicos con duraciones de 17 a 36 días, y magnitudes de M_w 5.4 a 6.2. El análisis de nuestros eventos utilizando leyes de escala momento-duración, revela valores consistentes con observaciones realizadas en otras zonas de subducción. El momento sísmico liberado por estos eventos es proporcional al cubo de su duración, lo que parece implicar una dinámica comparable con la de los terremotos clásicos. Los eventos detectados en este trabajo están principalmente localizados en zonas donde el acoplamiento intersísmico presenta valores en transición (0.3 - 0.8 de factor de acomplamiento), donde la zona de subducción transiciona de un estado bloqueado a uno de deslizamiento continuo. Finalmente, nosotros discutimos el rol potencial que podrían jugar los fluidos en el desencadenamiento de estos eventos lentos.

Résumé Depuis une vingtaine d'année, des événements de glissement asismiques ont été détectés le long de quasiment toutes les frontières de plaques au monde. Cependant, aucun n'a été décrit pour l'instant le long de la zone de Subduction allant du Perou au nord du Chili, si l'on omet le glissement mesuré lors de la période d'activité ayant mené au séisme d'Iquique en 2014. Nous utilisons une technique dite de Template matching sur des séries temporelles de déplacement mesuré par GNSS dans le nord du Chili pour extraire la signature d'événements de glissement asismiques cachés au sein du bruit géodésique. Nous détectons 24 événements asismiques avec des durées allant de 17 à 36 jours pour des magnitudes équivalentes allant de Mw 5.4 à 6.2. Nos événements ont des valeurs cohérentes avec les observations rapportées dans d'autres zones de subduction. Il apparait que ces événements asismiques sont essentiellement localisés dans des zones de couplage intermédiaires où le megathrust est a mi-chemin entre un état bloqué et un état en glissement permanent. Nous discutons finalement de l'influence éventuelle de fluides profonds dans le déclenchement de ces événements asismiques.

1 Introduction

Overwhelming evidence suggest that the Elastic Rebound Theory proposed by Reid (1910) after the 1906 California earthquake associated with the stick-slip behavior of frictional interface (Brace and Byerlee, 1966) is insufficient to explain the slip behavior along active faults. Geodetic measurements of surface motion have revealed the presence of aseismic, slow slip along all types of active faults. After the first descriptions in the mid-20th century from direct observations of damage to human-made structures crossing the San Andreas (Louderback, 1942; Steinbrugge et al., 1960) and North Anatolian (Ambraseys, 1970) faults, aseismic slip has been directly observed, or inferred, from geodetic measurements at different stages of the earthquake cycle. For instance, afterslip corresponds to the diffusion of slow slip during the post-seismic period accommodating a co-seismic stress perturbation (e.g., Heki et al., 1997; Bürgmann et al., 2001; Hsu et al., 2002, 2006). Creep, on the other hand, often refers to steady aseismic slip during the interseismic period (Steinbrugge et al., 1960; Ambraseys, 1970; Jolivet et al., 2015b). In addition, interseismic transients (i.e., slow slip events or SSEs) during this interseismic period were discovered in the 2000s along subduction zones. SSEs often locate in the deeper portion of the seismogenic zone (e.g., Hirose et al., 1999; Dragert et al., 2001), but some of these SSEs are associated with seismic signals that occur within the seismogenic zone, and may contribute to reducing geodetic coupling (Mazzotti et al., 2000; Bürgmann et al., 2005; Loveless and Meade, 2010; Radiguet et al., 2012; Béjar-Pizarro et al., 2013; Villegas-Lanza et al., 2016; Métois et al., 2016; Michel et al., 2019a; Jolivet et al., 2020; van Rijsingen et al., 2021; Lovery et al., 2024). This along-dip segmentation differs from one subduction zone to the other (Nishikawa et al., 2019) and we note more occurrences of SSEs along young, warm subduction zones (i.e., Nankai, Mexico, Cascadia), than old and cold ones. Finally, slow slip appears to be an important ingredient of the preparation phase of earthquakes (e.g., Ruegg et al., 2001; Ruiz et al., 2014; Radiguet et al., 2016; Socquet et al., 2017; Voss et al., 2018). More recently, it has been proposed that a significant fraction of observed geodetic displacement in seismically active regions results from the occurrence of slow slip events (Jolivet and Frank, 2020, and reference therein), suggesting a burst-like, episodic behavior of aseismic slip at all time scales from seconds to decades in places as varied as Mexico (Frank, 2016; Rousset et al., 2017; Frank and Brodsky, 2019), Cascadia (Michel et al., 2019a; Ducellier et al., 2022; Itoh et al., 2022), along the San Andreas Fault (Khoshmanesh and Shirzaei, 2018; Rousset et al., 2019; Michel et al., 2022), the Haiyuan fault in Tibet (Jolivet et al., 2015a; Li et al., 2021), on the Alto Tiberina and Pollino fault systems in Italy (Gualandi et al., 2017; Cheloni et al., 2017; Essing and Poli, 2022), or Japan (Nishimura et al., 2013; Takagi et al., 2019; Nishikawa et al., 2019; Uchida et al., 2020). All observations suggest the importance of accounting for

*Corresponding author: now at GFZ Potsdam, jorge@gfz-potsdam.de

Non-technical summary Earthquakes correspond to a sudden release of elastic energy stored in the crust as a response to the relative motion of tectonic plates. However, this release of energy is not always sudden and accompanied by destructive seismic waves. It sometimes happens slowly during aseismic, slow slip events. It has been shown that SSEs can be associated with the nucleation, propagation, and termination of big earthquakes. SSEs have been detected along many subduction zones in the world but not in northern Chile, yet. Here, we use a template matching method to scan GNSS observations of ground motion to detect and characterize slow slip events along the southern Peru - northern Chile subduction zone. We find 24 aseismic events at depths comparable with that of SSEs in other subduction zones, as well as in regions that slip aseismically persistently. We discuss how our findings relate to past earthquake ruptures, the geometry of the subduction zone, and fluids circulating at depth. Our results show the importance of implementing methods to extract small aseismic signals in noisy data, key observations for a better understanding of fault mechanics.

aseismic slip in our understanding of earthquake cycle dynamics. However, the underlying physics controlling aseismic slip is still debated, mainly due to the lack of good, dense observational databases.

Nowadays, observations of aseismic slip in subduction zones are frequently documented over a wide range of slip amplitudes and at different stages of the earthquake cycle (Avouac, 2015; Obara and Kato, 2016; Bürgmann, 2018; Kato and Ben-Zion, 2021, and references therein). Regular slow slip events have been documented mainly along warm subduction zones such as Cascadia, Nankai (southwest Japan), Mexico, or New Zealand (e.g., Graham et al., 2016; Nishikawa et al., 2019; Wallace, 2020; Michel et al., 2022, and references therein). Instead, observations of slow slip events in cold subduction zones such as off-shore Japan or Chile are sparse or indirect, through seismic swarms, repeaters, or slow earthquakes (Kato et al., 2012; Kato and Nakagawa, 2014; Gardonio et al., 2018; Nishikawa et al., 2019), and rarely with geodetic observations (Hino et al., 2014; Ruiz et al., 2014; Socquet et al., 2017; Boudin et al., 2021). Geodetic displacement corresponding to such slow slip events are usually of mm to cm-scale amplitude and require the development of novel and systematic methods to extract SSEs from noisy time series of geodetic data (Frank, 2016; Rousset et al., 2017; Michel et al., 2019a; Uchida et al., 2020; Itoh et al., 2022).

We focus on the South Peru- North Chile subduction zone. The region is seismically active, with two historical earthquakes in 1868 (southern Peru), and 1877 (northern Chile), both tsunamigenic earthquakes of magnitude ∼8.5 (Kausel, 1986; Comte and Pardo, 1991; Vigny and Klein, 2022) (Figure 1). Since these two events, the region has experienced several large earthquakes ($M_w > 7.5$) (Ruiz and Madariaga, 2018) accompanied by an important background seismic activity (Jara et al., 2017; Sippl et al., 2018, 2023) (Figure 1). In addition, coupling is highly variable along the subduction interface. Coupled regions overlap with the inferred rupture extent of the 2001 M_w 8.1 Arequipa and 2014 M_w 8.1 Iquique earthquakes (Schurr et al., 2014; Métois et al., 2016; Villegas-Lanza et al., 2016; Jolivet

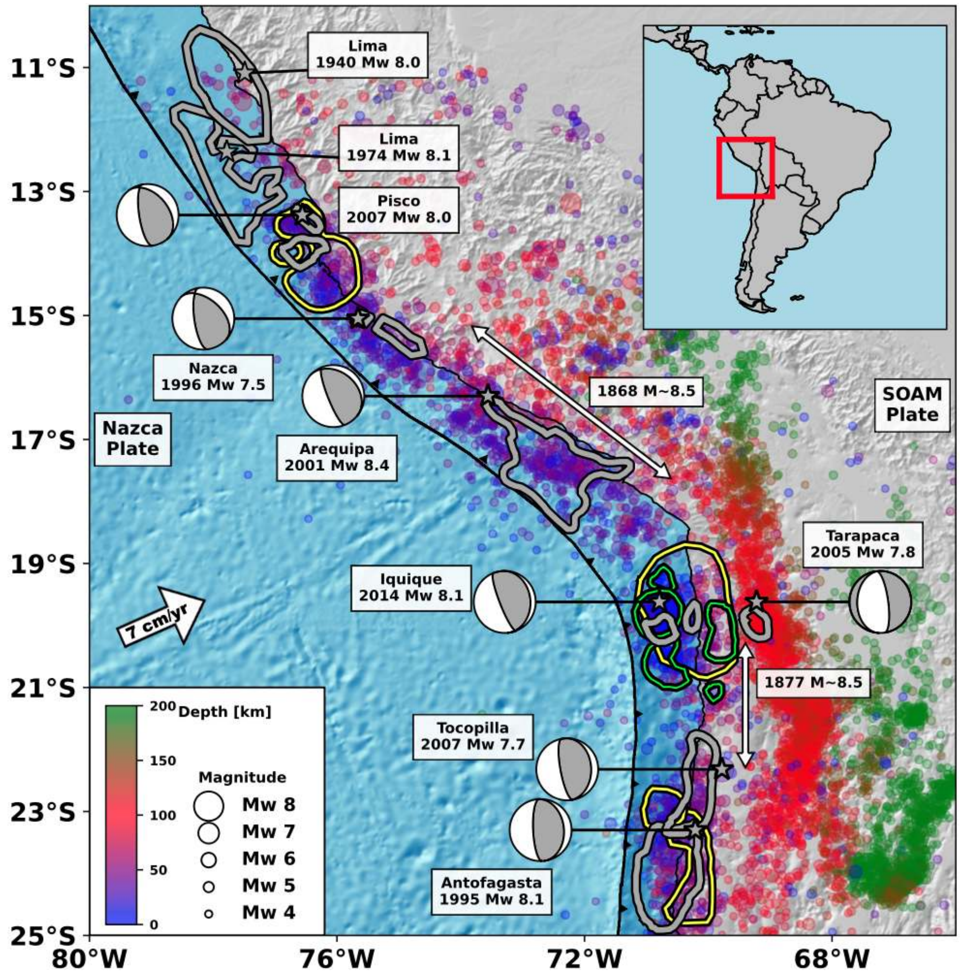


Figure 1 Seismotectonic map of the South Peru - North Chile subduction zone. White arrows show the extent of historical earthquakes (Comte and Pardo, 1991; Vigny and Klein, 2022). Gray contours are the rupture area of instrumental earthquakes with M>7.5, with corresponding epicenters (gray starts) and focal mechanisms (if available) (Dorbath et al., 1990; Beck and Ruff, 1989; Hartzell and Langer, 1993; Delouis et al., 1997; Chlieh et al., 2004; Pritchard et al., 2007; Dziewonski et al., 1981; Ekström et al., 2012; Peyrat and Favreau, 2010; Sladen et al., 2010; Béjar-Pizarro et al., 2010; Duputel et al., 2015; Jara et al., 2018). Yellow lines are the 0.1 m afterslip contours available in the region (Chlieh et al., 2004; Béjar-Pizarro et al., 2010; Remy et al., 2016; Hoffmann et al., 2018), whereas the green ones are the pre-seismic slip reported for Iquique earthquake by Socquet et al. (2017). Colored dots are earthquakes with M>4.0 from the International Seismological Centre (International Seismological Centre, 2016) over the period 1990 - 2016, color-coded by depth and scaled by magnitude. Large white arrow shows convergence direction and rate from Métois et al. (2016). SOAM: SOuth AMerica plate.

et al., 2020). A large coupled section is inferred where the 1877 earthquake is thought to have ruptured (Jolivet et al., 2020; Vigny and Klein, 2022). In addition, two low-coupling regions are observed. In southern Peru, low coupling coincides with the subduction of the Nazca ridge ($\sim 15^o$) (Villegas-Lanza et al., 2016; Lovery et al., 2024). In northern Chile, a reduction in coupling is inferred offshore Iquique and below the Mejillones peninsula ($\sim 21^o$) (Béjar-Pizarro et al., 2013; Métois et al., 2016; Jolivet et al., 2020).

In addition to low coupling, aseismic slip has been observed in South Peru and North Chile. Afterslip has been reported following large earthquakes, including the 1995 M_w 8.1 Antofagasta (Chlieh et al., 2004; Pritchard and Simons, 2006), the 2001 M_w 8.1 Arequipa (Ruegg et al., 2001; Melbourne, 2002), the 2007 M_w 8.0 Pisco (Perfettini et al., 2010; Remy et al., 2016), the 2007 M_w 7.7 Tocopilla (Béjar-Pizarro et al., 2010) and the 2014 M_w 8.1 Iquique earthquakes (Hoffmann et al., 2018) (Figure 1). Geodetic transients interpreted as the signature of aseismic slip occurred in the days to months preceding the M_w 8.4 Arequipa earthquake in 2001, before one of its largest aftershock, and preceding the Iquique earthquake in 2014 (e.g., Ruegg et al., 2001; Melbourne, 2002; Ruiz et al., 2014; Schurr et al., 2014; Socquet et al., 2017). Aseismic slip is considered responsible for a significant fraction of such geodetic transients (Twardzik et al., 2022). There is therefore plenty of evidence of occurrences of aseismic slip in this broad region but, despite intense efforts to instrument the area, no obvious spontaneous slow slip events have been detected during the interseismic period.

A change in the interseismic surface velocity field was observed following the M_w 7.5 intermediate-depth Tarapaca earthquake over a decade (Peyrat et al., 2006; Peyrat and Favreau, 2010) (Figure 1), an observation interpreted as the signature of a decoupling of the subduction interface (Ruiz et al., 2014; Jara et al., 2017). Comparable changes in surface velocity field, observed following the 2010 Maule earthquake, have also been observed in the regions affected by the 2015 Illapel (Ruiz et al., 2016) and 2016 Chiloé (Ruiz et al., 2017; Melnick et al., 2017) earthquakes. Such shifts in surface velocity may be linked to postseismic viscoelastic processes acting over long distances (Bouchon et al., 2018) in contrast to the localized behavior observed after the Tarapaca earthquake (Jara et al., 2017). Over the same period, we observed a significant increase in background seismicity (Jara et al., 2017), as well as an apparent synchronization of intermediate-depth and shallow seismic activities (Bouchon et al., 2016; Jara et al., 2017). Changes in background seismicity rates have been associated with the occurrence of aseismic slip events and fluid migration (Marsan et al., 2013; Reverso et al., 2016; Marsan et al., 2017). The synchronization of the seismicity is interpreted as related to aseismic slip events occurring along the subduction interface due to a broader slab deformation (Bouchon et al., 2016). These indirect observations suggest aseismic transients may occur in South Peru - North Chile during the interseismic period.

We aim to detect small, short-term aseismic slip events in this region and discuss their occurrence and location with respect to the interseismic coupling pattern and past seismic crises. We explore GNSS time series, searching for small transients, using a geodetic template matching approach (Rousset et al., 2017). We use GNSS and InSAR data to infer an updated distribution of interseismic coupling using a Bayesian framework following the approach of Jolivet et al. (2020), comparing the detected aseismic events with the coupling model, along with geophysical information available in the region (seismicity, Vp/Vs ratio, gravity models). We finally discuss potential mechanisms explaining the occurrence of aseismic events in the area.

2 Data, Methods and Results

2.1 GNSS processing and time series analysis

We process data from 119 continuous GNSS (cGNSS) sites in the central Andes region (Figure S1a) and worldwide (Figure S1b), using a double difference approach with the GAMIT/GLOBK software (Herring et al., 2015). 67 cGNSS sites are in the South Peru - North Chile region (Figure S1a and Figure 2, brown arrows), installed and maintained by the Integrate Plate boundary Observatory Chile (IPOC) (Klotz et al., 2017), the Laboratoire International Associé “Montessus de Ballore” (LIA-MB) (Klein et al., 2022), the Central Andean Tectonic Observatory (CAnTO, Caltech) (Simons et al., 2010), the Instituto Geofísico del Perú (IPG) (Jara et al., 2017; Socquet et al., 2017), the Institut des Sciences de la Terre (ISTerre) (Jara et al., 2017; Socquet et al., 2017), and the Centro Sismológico Nacional of Chile (CSN) (Báez et al., 2018). The remaining 52 stations are part of the International GNSS Service (IGS) (Teunissen and Montenbruck, 2017) global network. We separate these stations into three subnetworks (two locals and one global) with 33 overlapping stations, where the local separation depends on the station data span: one local network with data from 2000-2014 and the other including data from 2007-2014. Global network processing includes 99 stations over the 2000 - 2014 period, with 22 stations in South America (Figure S1b). We use the GAMIT 10.6 software (Herring et al., 2015), choosing the ionosphere-free combinations and fixing the ambiguities to integer values. We use precise orbits from the IGS, precise earth-orientation parameters (EOPs) from the International Earth Rotation and Reference System Service (IERS) bulletin B, IGS tables to describe the phase centers of the antennas, FES2004 ocean-tidal loading corrections, and atmospheric loading corrections (tidal and non-tidal). We estimate one tropospheric zenith delay every two hours and one pair of horizontal tropospheric gradients per 24h session using the Vienna Mapping Function (VMF1) (Boehm et al., 2006). We use the GLOBK software to combine daily solutions and the PYACS software (Nocquet, 2018) to derive position time series in the ITRF 2008 reference frame (Altamimi et al., 2011). Finally, time series are referenced to fixed South-America considering the Euler pole solution proposed by Nocquet et al. (2014).

We fit the time series with a parametric function of time for each component (N, E, and U) (Bevis and Brown, 2014). Each time series $x(t)$, function of time t, is modeled as

$$\begin{aligned} x(t) \;&= x_R \;+\; v(t-t_R) \;+\; \sum_{j=1}^{n_j} b_j H(t-t_j) \\ &+\; \sum_{k=1}^{n_F} [s_k \sin(\omega_k t) \;+\; c_k \cos(\omega_k t)] \\ &+\; \sum_{i=i}^{n_T} a_i \log(1 + t_i/\Delta T), \end{aligned} \tag{1}$$

where x_R is a reference position at a time t_R and v is the interseismic velocity for each component. H is a Heaviside function applied each time t_j an earthquake (or antenna change) offsets the time series. The combination of $\sin$ and $\cos$ functions describes seasonal oscillations (with annual and semi-annual periods), while the logarithmic function models the transient, post-seismic signal following large earthquakes ($M_w \geq$ 7.5) with a relaxation time ΔT. For a given station, we consider a Heaviside function for all earthquakes of magnitude larger that 6 with an epicenter to station distance lower than $\mathrm{d}(M) = 10^{\frac{M}{2}-0.8}$, as proposed by the Nevada Geodetic Laboratory (www.geodesy.unr.edu). We only include a post-seismic term for earthquakes of magnitude larger than 7.5. All inferred parameters for each component and each cGPS site are in Supplementary Information, Tables S1-S38. Figures S2-S17 compare the data and model at each station. We then estimate and remove a common-mode error by stacking all the time series (Bock and Melgar, 2016; Socquet et al., 2017; Jara et al., 2017). This procedure enables us to get residual time series (Figures S18-S19) as well as an interseismic velocity field (Table S1-S2). We use the obtained residual time series to search for geodetic transients compatible with slip on the megathrust and use the geodetic velocity field to update the last published coupling map (Jolivet et al., 2020).

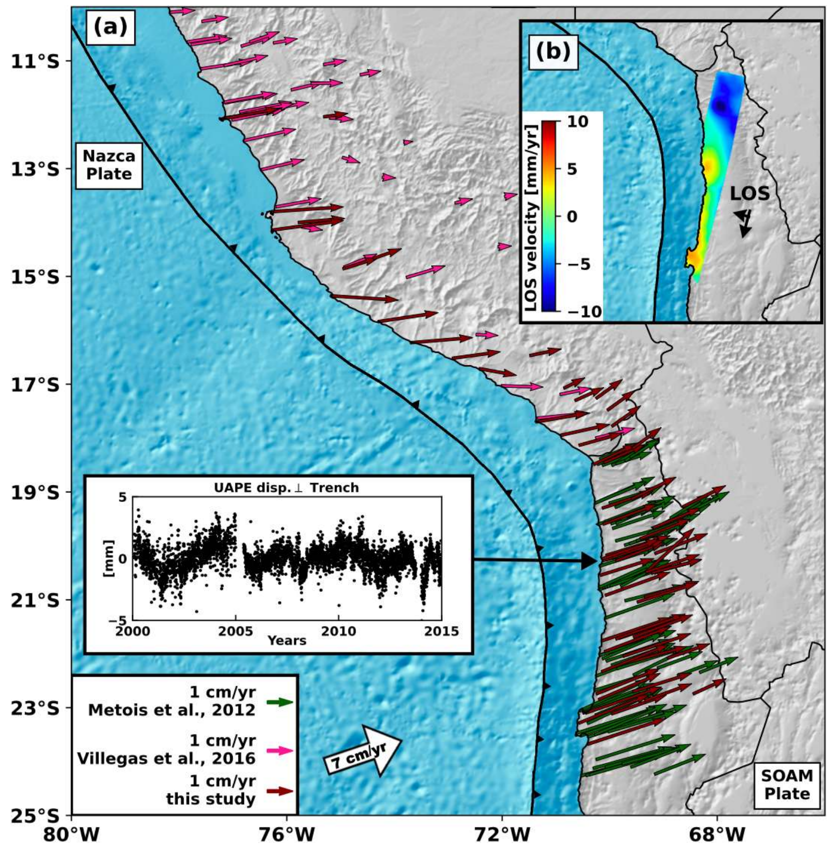


Figure 2 Geodetic data. (a) Colored dark green and pink arrows are the GNSS interseismic velocities from Métois et al. (2016) and Villegas-Lanza et al. (2016), respectively, while brown arrows are the continuous GNSS processed in this study. The inset shows the residual trench perpendicular displacement time series for GNSS station UAPE. (b) Line-of-sight (LOS) interseismic ground velocity from track 96 (Envisat data) from (Jolivet and Simons, 2018; Jolivet et al., 2020). Black arrows indicate the flight direction of the satellite and its line of sight (LOS).

2.2 Fault Geometry and Green's Functions

Coupling map estimation and geodetic template matching methods need a fault geometry and Green's functions calculation, as described below. In both cases, we define the geometry of the megathrust using Slab 2.0 (Hayes et al., 2018) as a reference, but with different meshing strategies. For the coupling case, we use triangles with 10 km-long sides along the coast and 25 km-long sides, both at the trench and depth, between latitudes 17^{o}S-25^{o}S. In the northern part (10^{o}S-17^{o}S), we adapt the size to the GNSS station density, considering a constant 50 km-long triangle side. In contrast, in the geodetic template matching case, we use triangles with 10 km-long sides along the coast and 25 km-long sides in the entire region. Then, we consider slip on the fault as the linear interpolation of slip values at the mesh nodes. Finally, we compute the Green's functions assuming a stratified elastic medium derived from Husen et al. (1999) using the EDKS software (Zhu and Rivera, 2002).

2.3 Coupling map for Southern Peru - Northern Chile

We update the distribution of coupling from Jolivet et al. (2020) in order to compare short- (i.e., days to months) and long-term (i.e., years to decades) aseismic deformation in the region. We use the GNSS velocity fields from Métois et al. (2016) (data span 1996 - 2013) and Villegas-Lanza et al. (2016) (data span 2008-2013), that we complement with our GNSS velocity field (Figure 2a, data span 2000-2016). Additionally, we use the line of sight (LOS) velocity map from Jolivet et al. (2020), derived from the processing of Envisat data covering the period 2003 - 2010 (Figure 2b).

We use the backslip approach to estimate the distribution of coupling (Savage, 1983). A coupling of 1 (resp. 0) corresponds to a fully locked megathrust (resp. a megathrust that slips at plate rate). We consider plate motion estimated by UNAVCO (www.unavco.org) under the ITRF 2014 model (Altamimi et al., 2016) to estimate the convergence rate, angle, and rake on each node of the fault mesh. The backslip rate is evaluated by subtracting the sliver movement proposed by Métois et al. (2016) in Chile (11 mm/yr) and by (Villegas-Lanza et al., 2016) in Peru (5.5 mm/yr) to the convergence rate. In the Arica bend (16^{o}S - 18^{o}S), at the boundary of the Chilean and Peruvian slivers, we build a gradient to make a smooth transition between the two slivers. We solve for the distribution of models that satisfy the geodetic data.

The forward problem is written as $\mathbf{d} = \mathbf{G}\mathbf{m}$, with $\mathbf{d}$ the geodetic data (GNSS and InSAR velocities), $\mathbf{m}$ the vector of parameters to solve for and $\mathbf{G}$ the Green's functions (Section 2.2). Parameters include coupling at each mesh node and geometric transformations akin to those in Jolivet et al. (2020). We adopt a probabilistic approach to estimate the parameters in order to evaluate the associated uncertainties. The *a posteriori* Probability Density Function (PDF) of a model $\mathbf{m}$ given a dataset $\mathbf{d}$, $\mathrm{p}(\mathbf{m}|\mathbf{d})$, writes as

$$\mathrm{p}(\mathbf{m}|\mathbf{d}) \ \propto \ \mathrm{p}(\mathbf{m})\mathrm{p}(\mathbf{d}|\mathbf{m}), \tag{2}$$

where $\mathrm{p}(\mathbf{m})$ is the *a priori* model PDF and $\mathrm{p}(\mathbf{d}|\mathbf{m})$ is the data likelihood. The *a priori* PDF describes our knowledge of coupling along the megathrust before collecting geodetic data. We define the *a priori* PDF at each node for the coupling factor as follows:

$$\mathbf{X} \sim \begin{cases} \mathcal{N}(\mu_c, \sigma_c^2) & \text{if } -0.1 \leq \mathbf{X} \leq 1.1 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where μ_c and σ_c are the mean and standard deviation of a normal distribution. We select the bounds of [-0.1, 1.1] to ensure an accurate sampling for the full range of coupling values between 0 and 1 (Dal Zilio et al., 2020a; Jolivet et al., 2020). We know the megathrust is decoupled below 60 km depth from geodetic (Chlieh et al., 2004; Béjar-Pizarro et al., 2013; Jolivet et al., 2020) and seismological evidence (Comte et al., 2016). Thus, we apply an *a priori* condition based on the depth of each node. If a node is deeper than 60 km, the *a priori* mean (μ_c) is set to 0 and the standard deviation (σ_c) to 0.1. In cases where a node is shallower than 60 km, we assign an *a priori* mean (μ_c) of 0.5 and a standard deviation (σ_c) of 0.5.

We adopt a Gaussian formulation for the data likelihood, $\mathrm{p}(\mathbf{d}|\mathbf{m})$, which writes as

$$\mathrm{p}(\mathbf{d}|\mathbf{m}) \ = \ \frac{1}{\sqrt{2}\mathbf{C}_\chi} \ \exp\left\{\left[-\frac{1}{2}(\mathbf{Gm}-\mathbf{d})^T \, \mathbf{C}_\chi^{-1}(\mathbf{Gm}-\mathbf{d})\right]\right\}, \tag{4}$$

where $\mathbf{C}_\chi$ is the misfit covariance matrix (Duputel et al., 2014) defined as $\mathbf{C}_\chi \ = \ \mathbf{C}_p \ + \mathbf{C}_d$, where $\mathbf{C}_d$ is the data covariance matrix (data uncertainties), while $\mathbf{C}_p$ is the prediction error covariance matrix, representing uncertainties on the assumed elastic model (P and S wave velocities and density). We assume a 10% error on the elastic parameters following Jolivet et al. (2020).

We explore the model space using Altar (altar.readthedocs.io) to sample the *a posteriori* PDF of the coupling factor, generating 250000 models. AlTar is based on the Cascading Adaptive Transitional Metropolis in Parallel (CATMIP) algorithm (Minson et al., 2013; Duputel et al., 2014; Jolivet et al., 2015b). These models enable us to perform statistics, derive the mean model for the interseismic coupling (Figure 3), and collect information about the model resolution (see Supporting Information for model GNSS and InSAR residuals, Figure S20-S23, as well as Standard Deviation, Mode, Skewness, and Kurtosis, Figure S24).

The mean coupling model (Figure 3a), is close to previously published models in the region (e.g., Chlieh et al., 2011; Béjar-Pizarro et al., 2013; Métois et al., 2016; Villegas-Lanza et al., 2016; Jolivet et al., 2020; Lovery et al., 2024), especially considering the along-strike segmentation. Our model differs from previously published models in the coupling intensity at locked patches, as well as the depth of these coupled patches. In Peru, we observe three patches with interseismic coupling that varies between 0.5-0.75 (Figure 3a). Previous models report similar patches, although totally locked (coupling factor ~ 1) (Chlieh et al., 2011; Villegas-Lanza et al., 2016; Lovery et al., 2024). Unfortunately, the density of GNSS stations in this region is not anywhere near that in Chile, hence the large standard deviations in the Peruvian region (Figure S25). Analyzing the moments of the *a posteriori* PDF, including standard deviation, skewness and kurtosis confirms this (Figure S24). Similarly, these moments show that the resolution at the trench over the entire region is low. Additionally, our model varies from those constrained only by GPS data in Chile (e.g., Métois et al., 2016). The InSAR data helps constraining interseismic coupling at depth (Béjar-Pizarro et al., 2013; Jolivet et al., 2020) and the strong *a priori* coupling damps potential large variations at depth, which we consider not physical.

2.4 Detection of aseismic slip events with geodetic template matching

2.4.1 Methodology

We use a geodetic template matching approach to detect potential aseismic slip events on the residual GNSS time series (Section 2.1). We summarize here the method presented in detail by Rousset et al. (2019). We search for the spatio-temporal signature of slip events in surface displacement time series by cross-correlating synthetic templates with our GNSS residual time series, in velocity. These templates correspond to the surface displacement caused by slip on dislocations located on the subduction megathrust embedded in a stratified, semi-infinite elastic medium. We calculate such templates (w) by convolving the Green's functions (Section 2.2) with a time-dependent slip evolution $\mathbf{s}(t)$ defined as

$$\mathbf{s}(t) \ = \ \frac{1}{2}\left[1 \ - \ \cos\left(\frac{\pi t}{T}\right)\right] \ , \tag{5}$$

where T is the duration of a synthetic event. Following Rousset et al. (2019), we derive for each template the weighted correlation function for each fault node, defined as

$$\mathbf{C}_f(t) = \frac{\sum_{i=1}^{2N} \mid \mathbf{G}_i \mid \mathbf{C}_i(t)}{\sum_{i=1}^{2N} \mid \mathbf{G}_i \mid} \ , \tag{6}$$

where $\mathbf{G}$ is the Green's functions and $\mathbf{C}_i$ is the correlation between the time series and the synthetic template at a given fault node i given by

$$\mathbf{C}_i(t) = \frac{\sum_{k=1}^{T} \dot{\mathbf{w}}_i(t_k)\dot{\mathbf{d}}_i(t_k + \tau)}{\sqrt{\sum_{k=1}^{T} \dot{\mathbf{w}}_i^2(t_k) \sum_{k=1}^{T} \dot{\mathbf{d}}_i^2(t_k + \tau)}} \ , \tag{7}$$

where $\dot{\mathbf{w}}$ and $\dot{\mathbf{d}}$ are the time derivatives of the template in terms of displacement (i.e., the template's velocity with duration T), and the time derivatives of the GNSS time

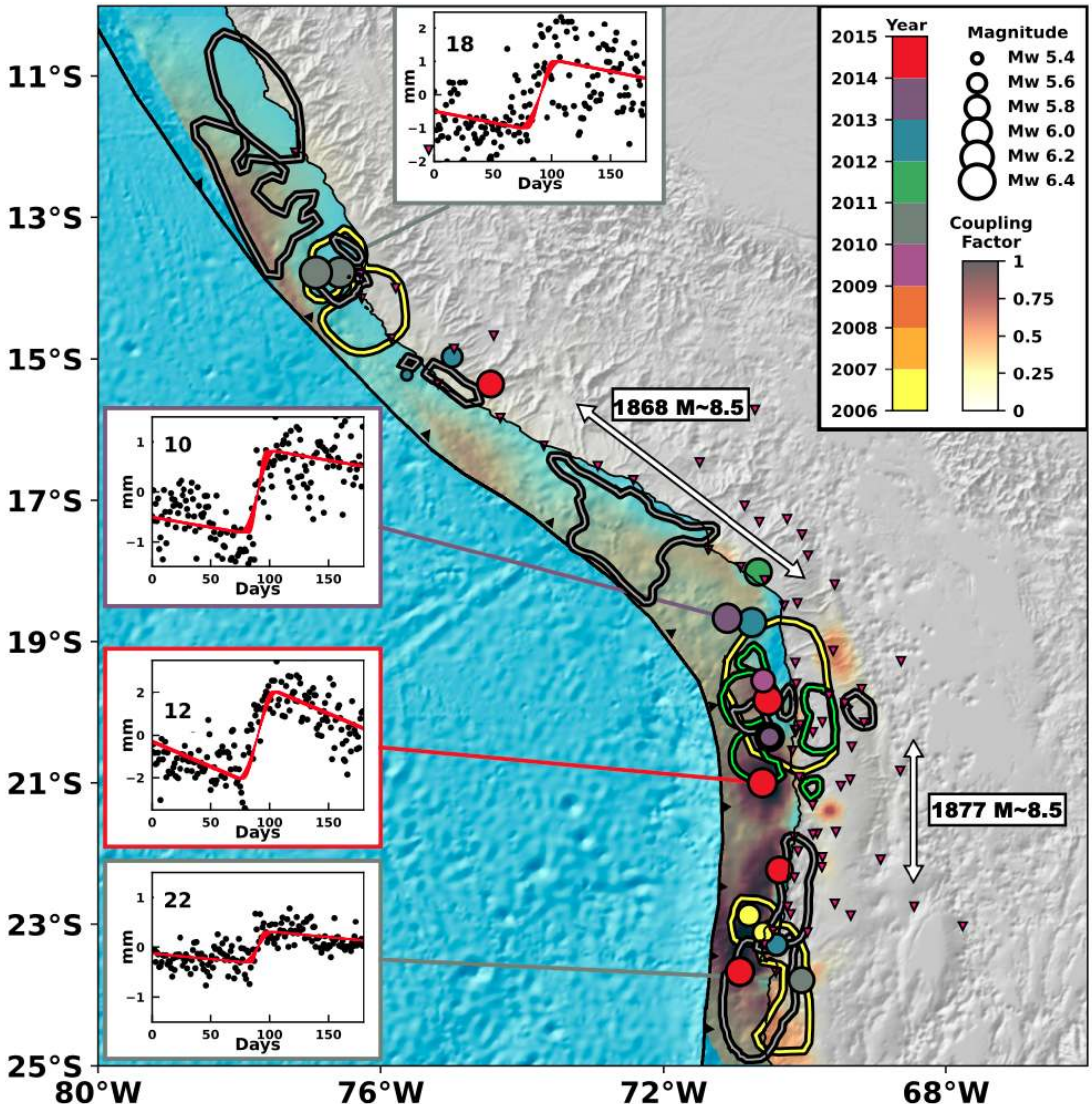


Figure 3 Location of detected aseismic slip events. Markers are color-coded by time of occurrence and scaled by magnitude. Four examples of weighted stacked correlations are shown with the event id number. Red line is the best fit model used to evaluate the event magnitude and duration, considering their estimated σ. Background color from white to dark through yellow and red is the mean coupling distribution. Black red areas (coupling factor $\sim$1) are locked regions, while transparent areas (coupling factor $\sim$0) are regions that slip aseismically at a rate equal to the plate convergence rate. Gray contours show instrumental ruptures. Yellow contours are afterslip regions, whereas green ones indicate slip inferred during the period preceding the Iquique earthquake. White arrows are the historical rupture extents.

series, respectively. τ denotes a moving time variable that enables the temporal matching search between templates and observations. We then search for peaks in $\mathbf{C}_f(t)$ corresponding to candidate slip events. As can be seen in the Supporting Information (see Fig. S31b, red and black lines), in the case of synthetic events, the correlation peaks in C_f arise from the geodetic noise using as many GNSS stations as possible.

For each candidate slip event, we stack the time series of displacement weighted by Green's functions around the time of detection (see Supporting Information Figure S31b, for an example of stacks on synthetic time series, purple and yellow lines). Such weighting accounts for displacement amplitude and direction, increasing the signal-to-noise ratio (Rousset et al., 2017). Stacks are computed over a period of 180 days, centered on each potential occurrence. On each stack, we estimate two linear trends, before and after the candidate occurrence, and the time dependent slip evolution of Eq. 5 to the weighted stack in order to determine the amplitude, the start and end date of each detected transient. We apply a non-linear regression to determine the posterior Probability Density Function of the model parameters given a stack of time series following Tarantola (2005). Effectively, we use an MCMC algorithm to derive 30,000 samples from the posterior PDF and evaluate the mean and standard deviation of the duration and magnitude of each candidate slow slip event.

In order to curate the potential detections from artefacts, we perform a sensitivity and resolution analysis, to determine the minimum magnitude of a slip event that can be detected for each fault node. Although the method above has been extensively described by Rousset et al. (2019), the novelty of our approach relies on the evaluation of uncertainties through a Bayesian exploration of all important parameters.

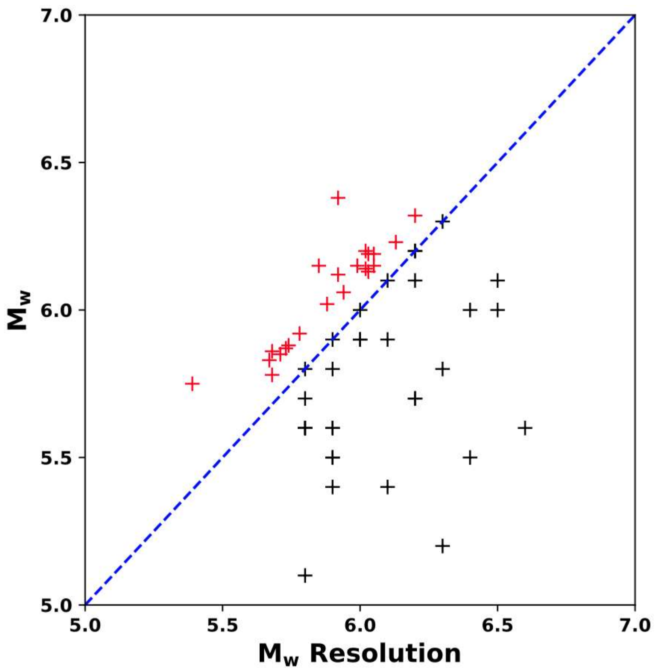


Figure 4 Event magnitude as a function of the resolution magnitude of the node where the event is located. Red crosses are events that passed the resolution test. Dashed blue line is the 1:1 line that separates validated from excluded events.

2.4.2 GNSS network sensitivity and resolution

We analyze the sensitivity of our approach by testing its ability to detect, locate, and estimate the source parameters (magnitude and duration) of synthetic aseismic slip events. We first evaluate the parameters characterizing the noise affecting each GNSS time series of displacement by building synthetic time series of noise on which we perform the tests. In order to generate synthetic noise, we model each component of the residual time series (Eq. 1) as a combination of white and colored noise (Williams, 2003), such as,

$$\mathbf{P}(f) \;=\; P_0\;(\mathbf{f}^{-\alpha} + f_0^{-\alpha}), \tag{8}$$

where $\mathbf{P}$ is the power spectrum as function of temporal frequency $\mathbf{f}$, P_0 and f_0 are normalization constants, and α is the spectral index. We explore P_0, f_0, and α using Bayesian inference to estimate their mean and standard deviation at each station component (see the Supporting Information for further details and an example of the power spectrum and the probability density function (PDF) of parameters at the UAPE station in Figures S26 - S27, as well as Tables S39 - S42 for all the network noise parameters inferred). We use these inferred noise parameters to build 1000 synthetic time series of displacement at each GNSS station. We use these synthetic time series to estimate thresholds of detection for each fault node.

The number of GNSS stations in the study area has evolved during the observation period. We, therefore, must consider three periods independently depending on the number of active stations: 2000 - 2003 (four stations), 2004 - 2007 (20 stations), and 2008 - 2014 (55 stations). We first determine which stations are able to capture a slow slip event on a given node. For each period and fault node, we correlate the 1000 synthetic time series of noise with a template of a duration of 40 days and slip equivalent to a magnitude M_w 6.0. We evaluate the standard deviation of the resulting weighted correlation functions, σ_t, as a minimal threshold to be exceeded (i.e., when dealing with time series that might include slip events, a peak of correlation higher than $3\sigma_t$ is a positive detection).

Once this threshold has been defined, we compute the weighted correlation function for 1000 time series of noise to which we have added the signal of synthetic transients with different duration (10, 20, and 30 days) and magnitudes (5.0 - 7.0 M_w, every 0.1 of magnitude). In case of a detection, we stack the displacement time series around the detection time. We consider a synthetic event has been correctly detected and located if we can recover four quantities, including the slip event location, timing, duration, and magnitude. If the estimated location is within 150 km from the true location, if the estimated timing and duration are within five days of the actual ones, and if the estimated magnitude is within 0.25 of the actual one, we consider the detection to be valid. This procedure enables us to determine the minimum magnitude that can be detected over each of the three observation periods and build resolution maps for each period investigated (see Supporting Information, Figures S29-S30). For instance, in the Iquique region ($\sim 19^o$S - 71^oW), the minimal magnitude M_w ranges from 6.6 to 6.8 from 2000 to 2003, decreases to 6.1-6.3 from 2004 to 2007 and again down to 5.9 to 6.1 from 2008 to 2014. Thus, as expected, we observe a significant improvement in detection sensitivity when the number of stations in a given region increases.

2.5 Application to GNSS time series

After exploring the network sensitivity to detect aseismic slip events, we search for transients in the residual time series obtained after subtracting the trajectory model described earlier. We fix the duration T of the template to 40 days and the slip to an event equivalent to M_w 6.0 (see Supporting Information, Figures S58-S59 for a test in the duration template sensitivity). By doing so, we detect 733 candidate slip events in the stacked correlation functions. Since some of these candidates may correspond to the same candidate slip event, we retain maximum occurrences within a radius of 150 km (i.e., if two maxima affect nodes separated by a distance higher than 150 km, they are considered as independent occurrences). After this selection step, we are left with 59 candidate slip events in the region. We evaluate their durations and magnitudes and compare these with our resolution maps. We keep candidates for which the obtained magnitude is higher than the minimum detectable magnitude for the corresponding node (Figure 4), leaving us with 24 validated slip events.

The duration of the slip events ranges from 17 to 36 days with magnitudes from M_w 5.4 to 6.2 and depths from 20 to 66 km. Figure 3 shows the location of the detected slip events along with four examples of weighted stacks. Figures 5 and 6 show two examples of stacks

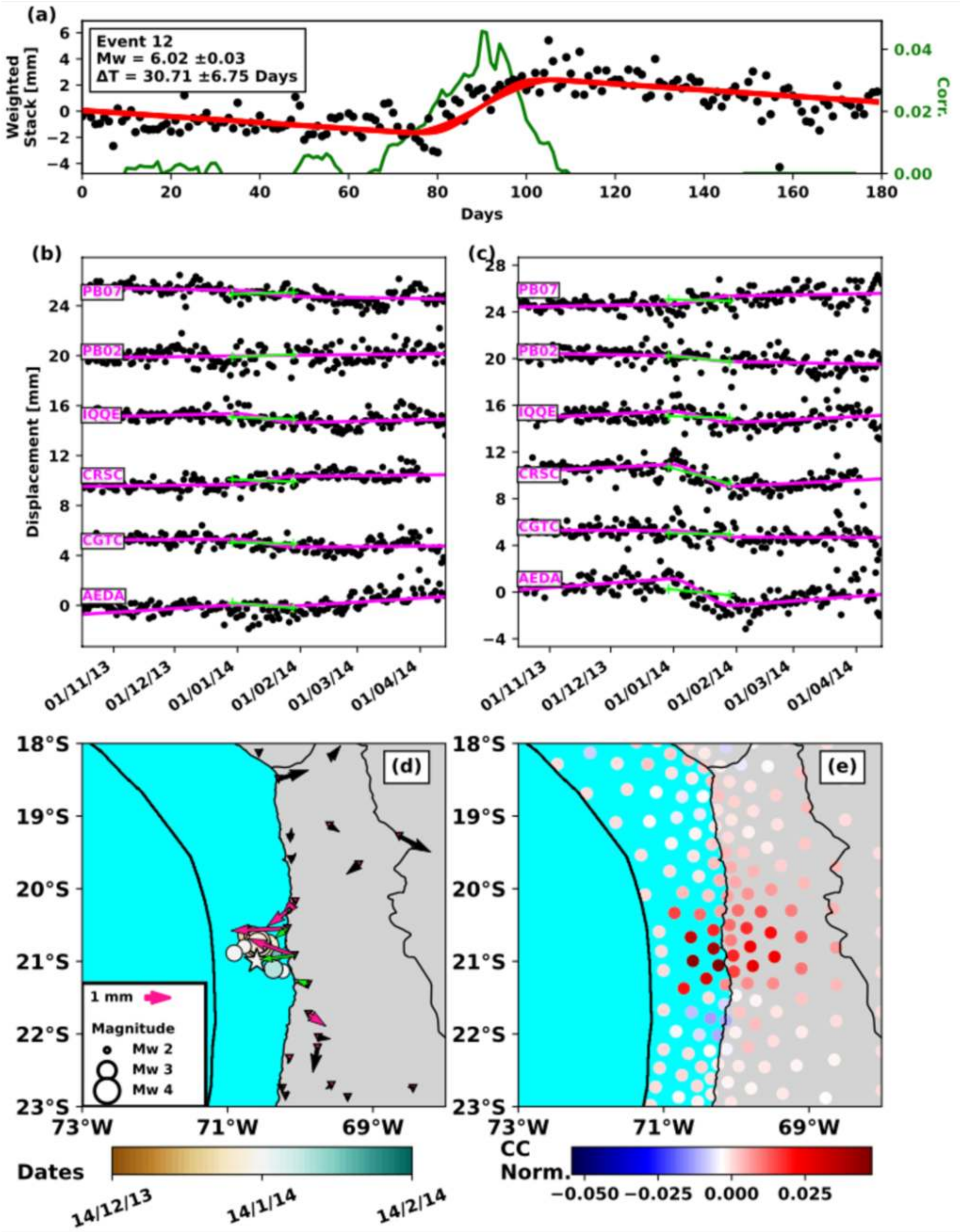


Figure 5 Example of detected aseismic slip event #12 in the vicinity of the 2014 Iquique earthquake, its locations, and associated seismicity. Figure (a) features the weighted stack for the event #12, with the red line representing the preferred model used to estimate event duration and magnitude, as indicated at the top left. The dark green line denotes the correlation function where event detection is made. Figures (b) and (c) display the displacement time series for the North and East components, respectively. Displacement data from six stations contributing to the weighted stack are shown. The pink lines indicate the best-fitting model for each displacement time series, which incorporates a linear trend and a transient, in accordance with Eq. 5. Meanwhile, the green lines represent the displacements for the estimated magnitude of each event. Figure (d) illustrates the envet location (marked by white star), with dots indicating seismicity before and after the event (spanning half of the event's duration for each period), scaled by magnitude and color-coded by date. Inverted triangles mark the GNSS station locations. Pink arrows denote the GNSS-derived displacements from observations used to estimate the weighted stack during the detected slow slip event, whereas black arrows indicate displacements not used in the estimation. The green arrows show displacements resulting from dislocations for the estimated magnitudes at each event location (white star). Figure (e) displays the map view of the correlation peak within the correlation function (illustrated in dark green in Figures a) for the event, pinpointing the moment when the detection is made.

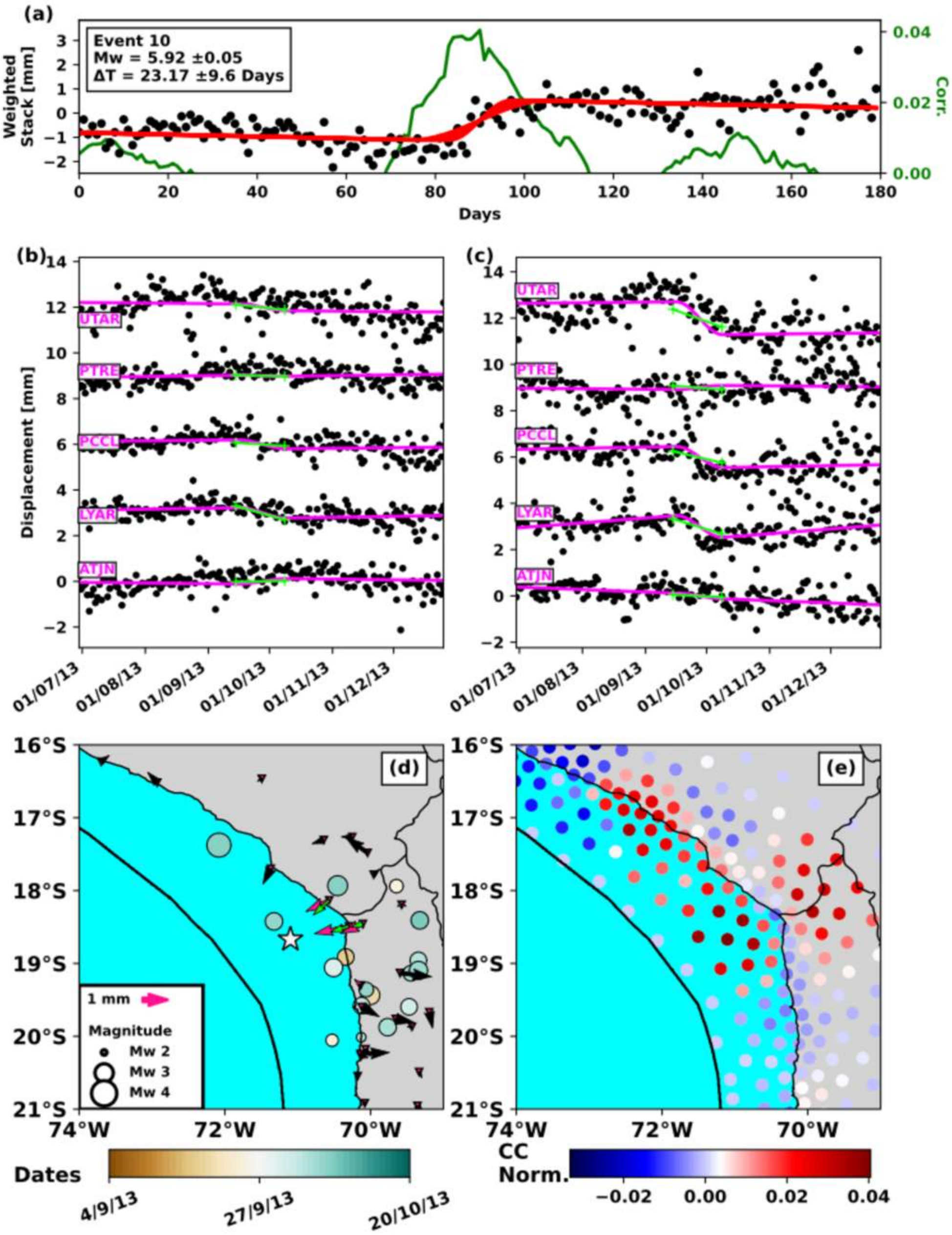


Figure 6 Same caption as Figure 5, but for event #10.

and correlation functions, along with the time series used to build the stacks and the map view of correlation peaks (see Supplementary information Tables S43 for the event parameters estimated with their uncertainties, and Figures S33 - S43 to see the data employed in the modeling, the data stack, and the model).

Following the methodology proposed by Nishimura et al. (2013), validated events are categorized into two types: probable and possible. This classification is achieved by comparing the displacement fields derived directly from observations with those generated by synthetics events of estimated magnitudes. Note that the magnitudes are estimated on the correlation stack and not directly on the measured displacements. A disagreement between the displacements corresponding to the detected magnitude on the detected node and the observed displacement would suggest our assumptions do not hold. Observed displacements are determined directly on the GNSS time series by estimating a linear trend along with a time-dependent slip evolution (Eq. 5). To estimate the displacement field for a detected magnitude, the slip corresponding to that magnitude is applied at the inferred location of each event. Figures 5 (b) and (c) illustrate examples of these estimates, with the actual displacements shown in magenta, while the displacements predicted from the magnitudes of each event are shown in green for Events #10 and #12 (see Supplementary Information, Figures S33 - S43 for the rest of the events). Upon analysis, we find that the agreement between observed and modeled ground motion is acceptable for 10 of our events, leading us to classify these as probable (A events, Table S43). Meanwhile, we observe a weaker agreement for 14 events which we hence categorize as possible (B events, Table S43).

Since our template matching approach only considers GNSS observations, we must ensure that the detected slip events (A and B) are mostly aseismic. We cross-check the 24 positive detections with the seismic catalog provided by the ISC (International Seismological Centre, 2016). We randomly generate 10000 synthetic locations for each slip event considering a normally distributed location uncertainty based on our resolution tests and estimate the sum of the seismic moment of all earthquakes occurring within at least a 2-σ radius of the detected slip event. We then compare this estimate of the seismic moment to the estimated aseismic one. All the detected slip events have an equivalent magnitude at least twice larger than the seismic magnitude (aseismic/seismic ratio for each event and further details on ratio estimation are in Supplementary Information, Table S43). Figures 5 and 6 (d) present the location of the two events detailed in Figures 5 and 6 (a) together with the seismicity that coincides with the occurrence of the slip event. These two events occur during the preparation phase of the 2014 Iquique earthquake (Event #12, Figure 1) and during the interseismic phase (Event #10). The combination of synthetic tests and the seismic vs. aseismic moment analysis confirms we detected 24 aseismic slip events (A and B) along southern Peru - northern Chile subduction zone over the period 2006 - 2014.

3 Discussion

3.1 Aseismic slip events and scaling laws

Aseismic slip events are now frequently observed along most subduction zones in the world, but the underlying physics is still debated. Among the points of debate, the comparison between slow slip and earthquakes should allow to point out whether comparable physics are involved. Ide et al. (2007) have proposed that, while the seismic moment of earthquakes is proportional to the cube of their duration, the moment of slow earthquakes, from tremors and low-frequency earthquakes to slow slip events, is proportional to the duration. Considering that simple considerations about size and stress drop led to the emergence of the observed scaling for earthquakes, the difference in moment-duration scaling should involve a fundamental difference between the mechanics of slow slip and that of earthquakes. Peng and Gomberg (2010) argued that the apparent moment duration scaling of slow earthquakes proposed by Ide et al. (2007) was only due to a lack of observations, suggesting both rapid and slow slip were driven by the same mechanism, namely a slip instability with variable speed and stress drop propagating along a weakened fault surface. In addition, Gomberg et al. (2016) proposed that seismic moment scales either with the duration or the cube of the duration depending on whether the rupture was elongated and pulse-like or mostly crack-like. Michel et al. (2019b) confirmed that the moment of slow slip events in Cascadia scales with the cube of their duration although being elongated and pulse-like. These observations agree with recent studies of aseismic slip and tremors in Japan (Takagi et al., 2019; Supino et al., 2020) and Mexico (Frank and Brodsky, 2019), as well as numerical modeling using dynamic simulations of frictional sliding (Dal Zilio et al., 2020b). Such numerical and observational evidence suggests that SSEs might exhibit comparable scaling as classical earthquakes, only with lower rupture speeds and stress drops.

We evaluate the scaling between moment and duration for the aseismic slip events we have detected. We estimate that the moment, M, is such as $M \propto T^{4.99\pm0.48}$, with T the duration for the 24 detected SSEs (refer to Figures 7, S45, and S46 in the Supporting Information for an in-depth explanation of the scaling estimation procedure). This scaling relationship remains consistent when analyzing events A ($M \propto T^{5.05\pm0.59}$, see Figures S47 and S48) and B ($M \propto T^{4.89\pm0.52}$, illustrated in Figures S49 and S50) independently. Our events seem to align with a moment-duration scaling T^3. However, as extensively discussed by Ide and Beroza (2023), uncertainties associated with the estimation of event duration might influence significantly our results. Consequently, it is challenging to definitively conclude that our findings adhere to the moment-duration T^3 scaling. That said, our detections are situated within the range of moment-duration observed in other subduction zones such as Cascadia, Japan, or Mexico (Ide and Beroza, 2023, and references therein). Building on this observation, we adopt the methodology outlined by Gomberg et al. (2016) to deduce the source properties of our events. We infer that the rupture velocities of our detections range between 0.5 and 10 km/day, accompanied by a stress drop of 0.1 MPa (see the Supporting Information for detailed information on the parameter estimation process). Although our method does not allow to detect events that would propagate, we observe our SSEs are more compatible with crack-like, unbounded ruptures than pulse-like, bounded ones. As a conclusion, our findings along southern Peru - northern Chile region align with SSEs observations from other subduction zones.

3.2 Aseismic slip and interseismic coupling distribution

Our coupling estimate corresponds to an average behavior over a decade, without accounting for potential slow slip events hidden within the noise. The slow slip events we detect hence correspond to fluctuations around this average. We compare the map of coupling to the location of our 24 aseismic events to explore how such fluctuations distribute with respect to locked and creeping asperities along the megathrust (Figure 3). We compare the distribution of coupling where our events are located to a distribution coupling at randomly picked locations (Figure 8, see the Supporting Information for a detailed explanation of the calculation of the PDF for coupling and detected events). The distributions differ but mostly when considering only events in northern Chile, where our estimate of coupling is much more robust. Detected slow slip events occur mostly in regions of intermediate coupling. This observation is not as clear for the Peruvian region, probably because of the sparsity of the data used here, although the same tendency is suggested on Figure 8. This result aligns with

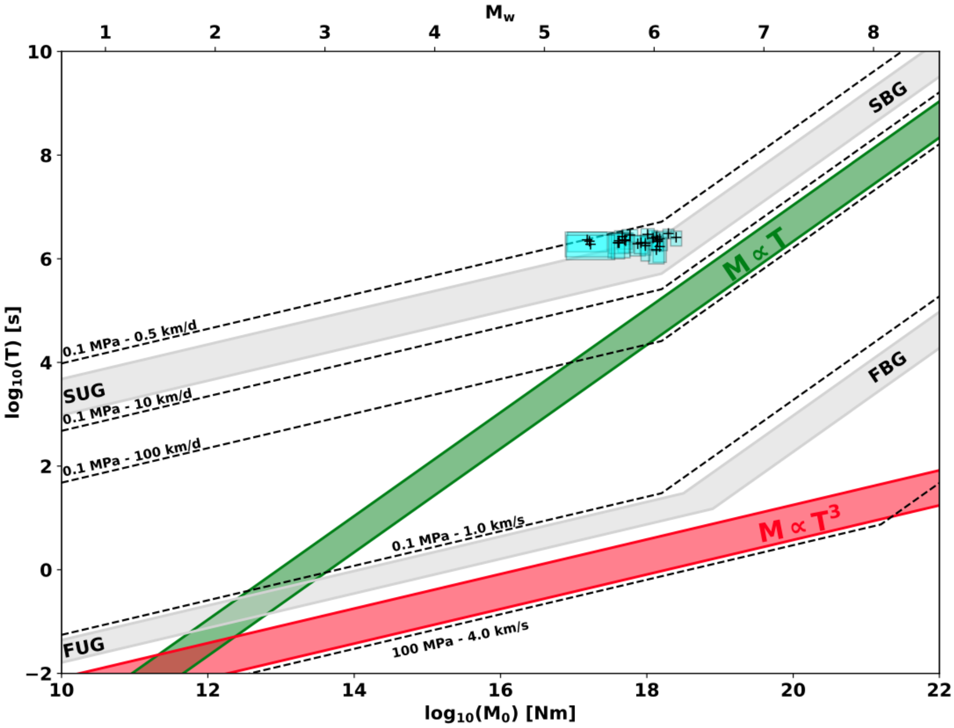


Figure 7 Seismic moment versus duration for our aseismic slip events following the scaling law proposed by Gomberg et al. (2016). Slow bounded/unbounded (SBG, SUG) and fast bounded/unbounded (FBG, FUG) regions are shown by light gray areas. Dashed lines are the theoretical relationship between moment and duration for a few selected stress-drop and rupture velocity values. The $M \propto T$ scaling is shown in green. The $M \propto T^3$ scaling is shown in red.

Frank (2016) findings in the Mexico subduction zone, where a database of slow slip events seems to compensate the lack of slip deficit in transition zones with respect to coupled regions of the megathrust. Materna et al. (2019) describe a comparable behavior over longer periods where coupling variations seem to occur in regions of transitional coupling (Michel et al., 2019a). In addition, events offshore Peru tend to cluster spatially around locked asperities, areas that are generally of intermediate coupling (Figure 9). In general, slow slip events occur in transitional regions between seismic asperities and freely slipping areas. This is consistent with model predictions from rate-and-state friction in which slow slip events are expected to occur at the transition between seismic, rate-weakening and creeping, rate-strengthening asperities (e.g., Liu and Rice, 2005, 2007; Perfettini and Ampuero, 2008).

The average depth of the detected slow slip events is 33 km (Figure 8, see the Supporting Information for a detailed explanation of the PDF calculation). Separating the events, by region, yields an average depth of 37 km for Peru and 30 km for northern Chile with comparable standard deviations (19 and 10 km respectively, Figure 8). This result remains consistent when conducting separate analyses of events A and B (refer to Figures S56-S57 in the Supporting Information). Lay (2015) separates the subduction megathrust along depth into four domains (A, B, C, and D). Domain A, located between the trench and a depth of about 15 km, hosts either tsunami earthquakes or aseismic deformation. Domain B, between approximately 15 and 30 km depth, hosts large megathrust earthquakes. Domain C, between approximately 30 and 50 km depth, hosts intermediate sized earthquakes. At greater depths, Domain D, between 50 and 70 km, hosts slow slip events, tremors, and very low-frequency earthquakes. Our slow slip events mainly occur in Domains C and D. It is understood that small, velocity weakening asperities in Domain C are embedded in conditionally stable regions of the megathrust, prone to host slow slip events. Domain D is dominated by aseismic sliding and potential slip rate variations could explain deeper detections. Therefore, the depth distribution of our events matches regions where slow slip events are expected in a subduction zone context.

Our resolution tests (Figures S24, S29-S30) suggest that it is impossible to capture aseismic slip near the trench, in domain A, with the current GNSS network. However, large, shallow slow slip events have been observed in Japan (Nishimura, 2014; Nishikawa et al., 2019) and New Zealand (Wallace, 2020). Seafloor geodesy might help to detect the occurrence of such large events and potentially for small, cm-scale ones comparable to our aseismic slip events (Araki et al., 2017). Additionally, stress-shadow induces apparent coupling in velocity-weakening regions, especially late in the interseismic period (Hetland and Simons, 2010; Lindsey et al., 2021). For this reason, we also cannot rule out the potential occurrence of aseismic slip event near the trench.

In addition to the depth-dependent segmentation, we observe an along-strike segmentation in the distribution of SSEs. In particular, we observe a lack of events within the rupture area of the 1877 earthquake, within the Arequipa rupture area and other detections gather around locked asperities, like in the doughnut model

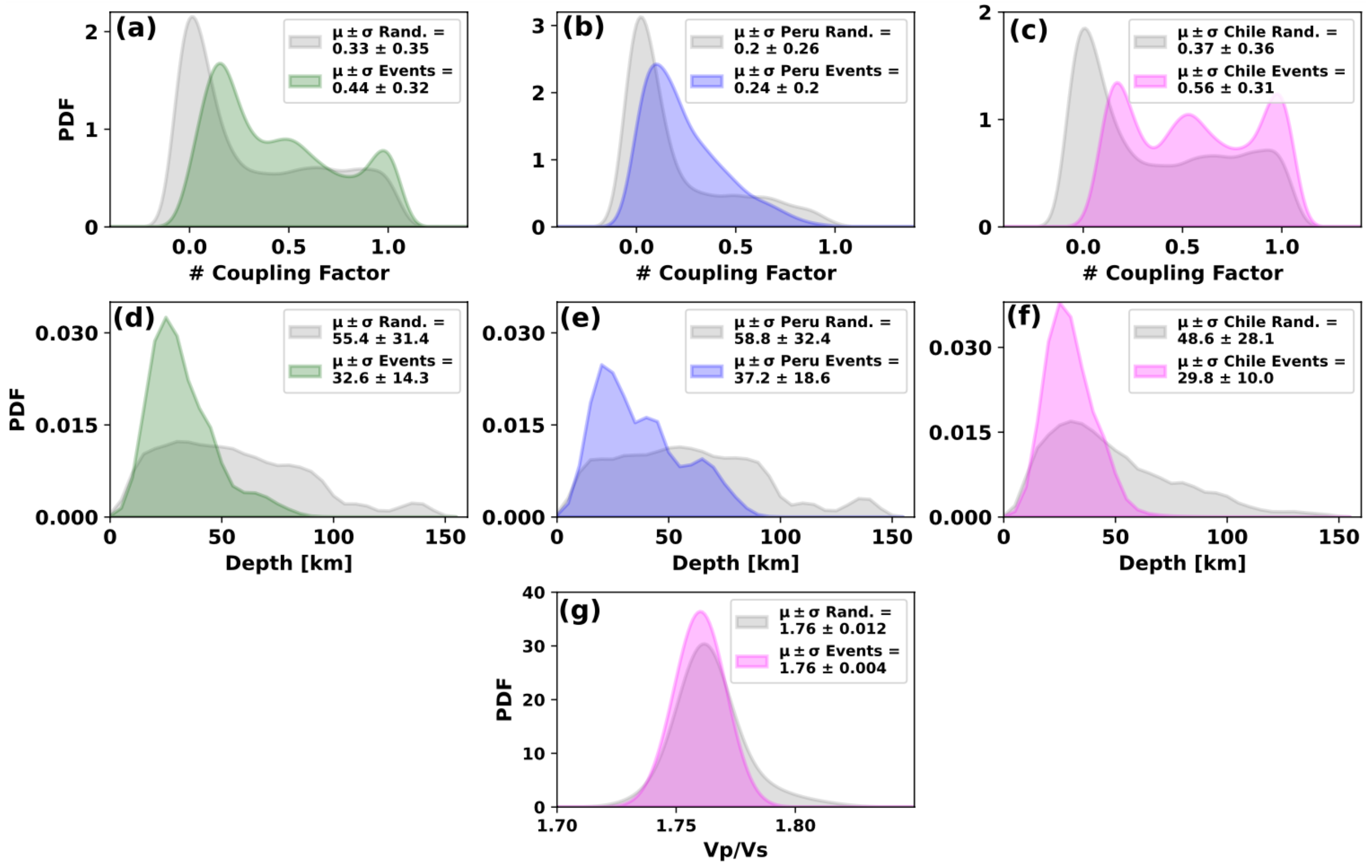


Figure 8 Coupling, depth, and Vp/Vs ratio of the detected aseismic slip events. (a) Probability Density Functions (PDF) of 1000 coupling models for 24 random picks (gray) and PDF of coupling where 24 aseismic slip events are detected (green), with respective mean (μ) and standard deviations (σ). (b) and (c) are the same as (a) for the Peru region only (gray: random, blue: SSEs) and northern Chile only(gray: random, magenta: events), respectively. (d) PDF of the depths of 24 random events (gray) and aseismic slip events detected in the region (green). (e) and (f) Same as (d) but for Peru (gray: random, blue: events) and Chile (gray, magenta) regions. (e) PDF of the Vp/Vs ratio for the Chilean region (gray, 17 random events), and detected aseismic events in Chile (magenta).

for seismicity (Kanamori, 1981; Schurr et al., 2020). Such configuration is comparable to that of the Japan trench where the asperity that ruptured during the Tohoku earthquake in 2011 overwhelms the simple depth-dependent distribution of behavior from Lay (2015). In particular, Nishikawa et al. (2019) propose that, unlike the Nankai subduction interface which exhibits a depth-dependent segmentation due to a young, warm slab, the megathrust beneath Tohoku is not segmented at depth into four distinct domains. In our area of interest, the subducting slab is older than the Nankai slab and probably colder (Müller et al., 2008), which would explain why the behavior we unravel is not completely consistent with that of Lay (2015) and potentially closer to that of the Japan trench.

As an additional level of complexity, three events coincide with the subduction of the Nazca ridge (14^{o}S, Figures 3 and 9a), six events are located beneath the Mejillones Peninsula (23^{o}S, Figures 3 and 9d), and three events are within the Arica bend (17^{o}S - 19^{o}S, Figures 3 and 9b and c). These morphological structures are anomalies compared to the model proposed by Lay (2015) as they are considered as barriers to the propagation of large earthquakes (Armijo and Thiele, 1990; Comte and Pardo, 1991; Béjar-Pizarro et al., 2010; Villegas-Lanza et al., 2016; Poli et al., 2017). In these regions, the depth of our detected slow slip events does not match the depth-dependency described by Lay (2015). We can speculate that local geometrical complexities may lead to the occurrence of slow slip events (Romanet et al., 2018) in the case of the subduction of the Nazca Ridge or that the apparent low coupling is the result of multiple slow slip events (Jolivet et al., 2020) in the case of the Arica Bend.

3.3 Aseismic slip events before and after large earthquakes

Among all the detected slow slip events, only events #7, and #12 (Figure 3, S36 and 5) do not occur during the steady interseismic period. Event #7 locates in the region struck by the Iquique earthquake in 2014 (Figure 9c, and S36) during the post-seismic relaxation that followed the mainshock (Meng et al., 2015; Hoffmann et al., 2018; Shrivastava et al., 2019) (M_w 6.1 and duration of 28 days in June 2014). Such slow slip events embedded within a post-seismic sequence have already been observed following the Illapel earthquake (Tissandier et al., 2023) and in a completely different setting, following the 2004 Parkfield earthquake, along the San Andreas Fault (Michel et al., 2022).

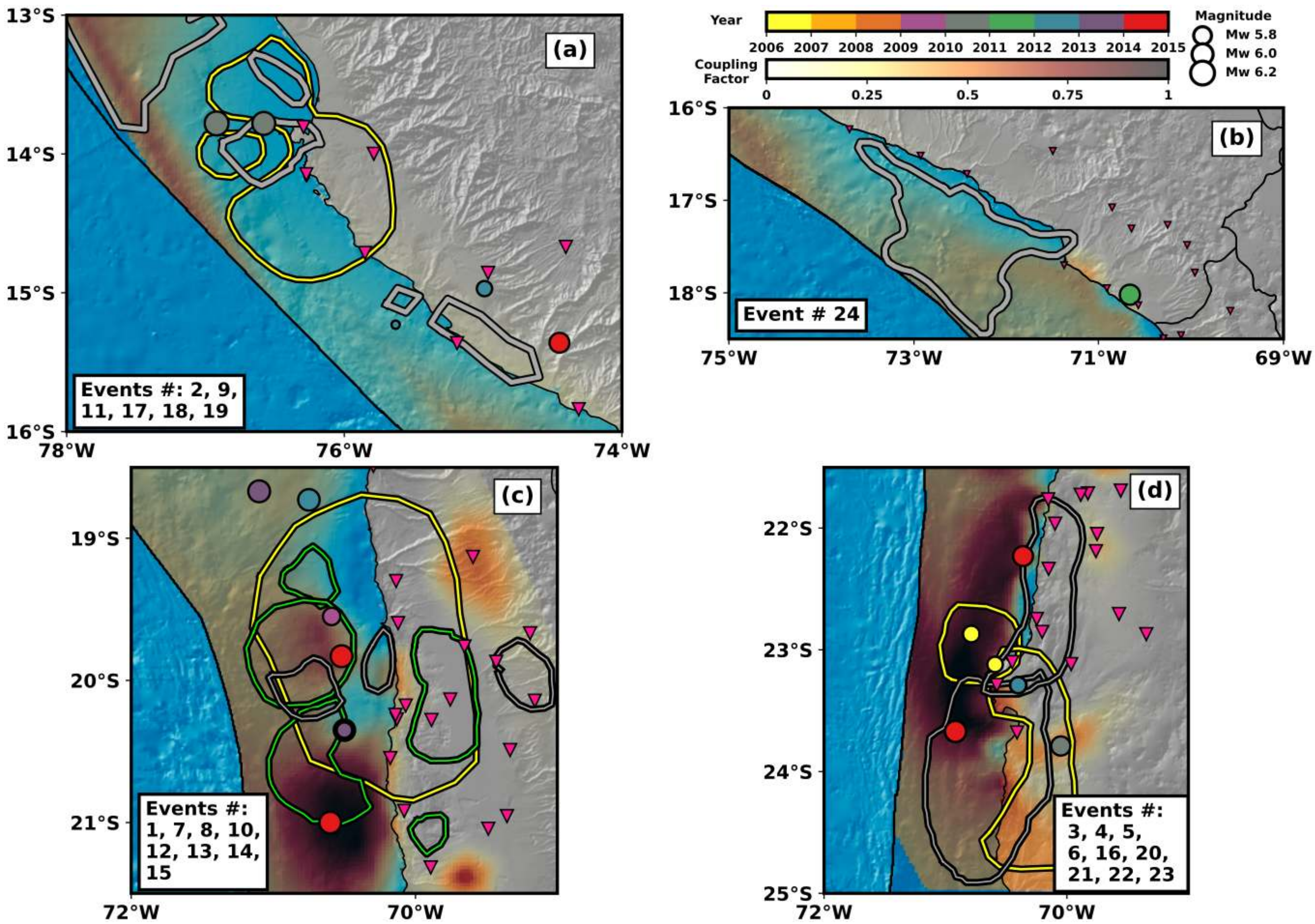


Figure 9 Zoom over a selection of regions of interest. Gray contours are instrumental ruptures. Yellow contours show reported afterslip. Our aseismic slip events are color-coded by time and scaled by magnitude. Background color shows our Bayesian inference of coupling. Inverted pink triangles are the GNSS stations used in this study. (a) Region struck by the Pisco (2007) and Nazca (1996) earthquakes. Our detections seem to cluster around asperities broken during earthquakes or afterslip regions. (b) Region struck by the Arequipa (2001) earthquake. (d) Region struck by the Iquique earthquake in 2014. Green contours show the preseismic slip reported by Socquet et al. (2017). Events occur around locked interseismic patches or low-coupled regions. (d) Region struck by the Antofagasta (1995) and Tocopilla (2007) earthquakes. Events surround broken asperities or locked interseismic patches, with a cluster beneath Mejillones Peninsula, potentially associated with earthquake afterslip. For citations of instrumental ruptures and afterslip, please refer to Figure 1

Aseismic slip has been recognized as an important element of the earthquake preparation phase (Obara and Kato, 2016; McLaskey, 2019; Kato and Ben-Zion, 2021, and references therein). An 8-month-long slow slip event was reported before the Iquique earthquake in 2014 (Socquet et al., 2017), and event #12 coincides with one of the regions of the megathrust that slipped aseismically during that preparation phase (Figure 9 c). In addition, event #12 occurred where and when intermediate-depth and shallow seismicity synchronized before the Iquique earthquake (Bouchon et al., 2016; Jara et al., 2017) (M_w 6.0 and duration of 30 days in January 2014). Such synchronization of seismicity began in January 2014, lasted for one month, and is interpreted as evidence of a slow, slab-wide deformation process prior to megathrust earthquakes (Bouchon et al., 2016). Furthermore, event #12 is coincident with the transient event reported by Boudin et al. (2021) using a long-base tiltmeter. Our epicentral location differs by ∼50km from the one reported by (Boudin et al., 2021), a difference that can be explained by different modeling strategies and/or uncertainties. We propose that event #12 is linked to the 8-month aseismic slip transient observed preceding the 2014 Iquique earthquake. Such detection suggests the growing instability preceding the Iquique earthquake exhibits a complex spatio-temporal behavior that hides within the noise of the data, in agreement with the hypothesis proposed by Jolivet and Frank (2020) and Twardzik et al. (2022).

3.4 Aseismic slip and fluids

Fluids may also play a role in the occurrence of aseismic slip events (Avouac, 2015; Harris, 2017; Jolivet and Frank, 2020, and references therein). Pore pressure affects fault normal stress, hence modify the probability of a slip instability as well as the nucleation size (Liu and Rice, 2007; Avouac, 2015; Bayart et al., 2016; Harris, 2017; Bürgmann, 2018; Jolivet and Frank, 2020; Behr and Bürgmann, 2021). An increase in pore pressure within the fault zone leads to a decrease in normal

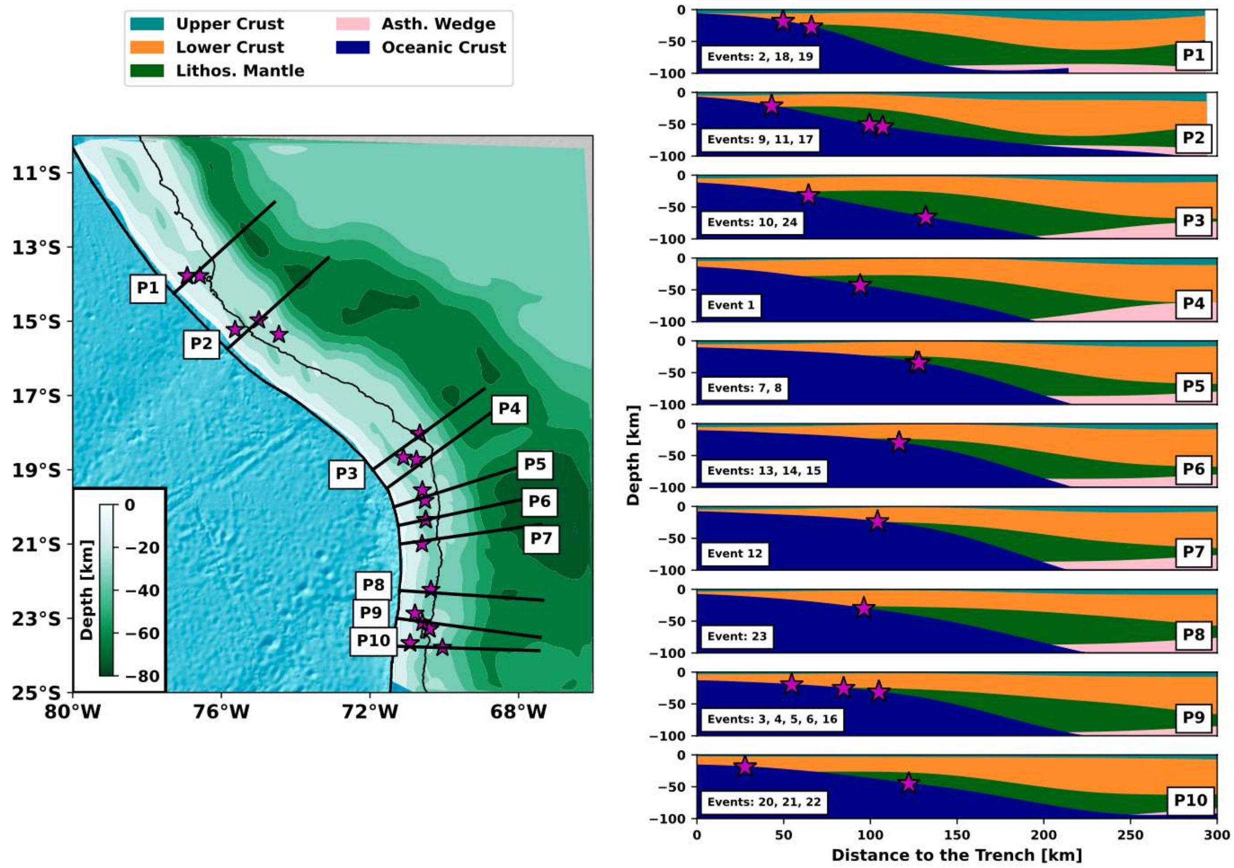


Figure 10 Map view of the depth of the continental Moho discontinuity from gravity-derived structural models by Tassara and Echaurren (2012). Magenta stars are the location of our 24 aseismic events. Black lines indicate the location of the profiles shown on the right. Colors indicate the structure at depth (upper and lower crusts, lithospheric mantle, asthenospheric wedge, and oceanic crust). White box indicates the id of events occurring along each profile.

stress, which promotes slip but increases nucleation size, promoting slow slip. We compare our detections to the distribution of the Vp/Vs ratio and to gravity-inferred structural models in the region. We use the Vp/Vs ratio inferred by Comte et al. (2016) for the events located in Northern Chile. Statistically, the 17 aseismic events in northern Chile are not related to a specific Vp/Vs value (Figure 8, see the Supporting Information for a detailed explanation of the PDF calculation). In particular, no slow slip events are found to collocate with high Vp/Vs ratios (Vp/Vs > 1.8) (Comte et al., 2016) (Figure S44).

We also compare the location of our aseismic events to a 3-D density model in the region (Tassara and Echaurren, 2012). Figure 10 shows the location of aseismic events along ten different trench-perpendicular cross sections. The slow slip events are primarily located along the contact between the slab and the overriding lithospheric mantle (Figure 10, see Figure S51 for an analysis of depth uncertainties). This mantle corner is principally hydrated by the dehydration of the subducting slab due to water releasing metamorphic reactions (Peacock, 2001; Rüpke et al., 2004; Comte et al., 2016; Wang et al., 2019; Contreras-Reyes et al., 2021). The fact that our aseismic slip events tend to cluster at depths corresponding to the lithospheric mantle along the megathrust, and not deeper, might imply that fluids may be trapped and accumulate below the continental Moho, an hypothesis that would require further investigations.

4 Conclusions

We have systematically analyzed GNSS time series in the region, searching for the occurrence of aseismic slip events with a template matching approach. We find 24 events in the period 2006 - 2014, with durations of 17 - 36 days, magnitudes of M_w 5.4 - 6.2, and located at depths of 20-66 km. These events are mostly aseismic and are observed at all stages of the earthquake cycle, including during post-seismic periods (afterslip, one event), earthquake preparation phase (one event), and interseismic period (22 events). We compare those slow slip occurrence to a wide range of possible models of interseismic coupling based on GNSS and InSAR velocity fields and infer a distribution of coupling along the megathrust.

By conducting a moment-duration scaling analysis, we find that our observations are consistent with values reported in subduction zones globally. We do not find particular correlations with published seismic velocity structures but find that slow slip events cluster around past ruptures and locked asperities, where the megathrust transitions from sliding to locked. Additionally, our events are located in regions of intermediate coupling values and mean depths of 33 km, which match regions where slow slip events occur in the context of subduction zones.

Some of these events occur on the subduction interface deeper than than the continental MOHO, i.e. where the slab is in contact with the mantle wedge corner where fluids are supposedly trapped. This points toward the influence of fluids as it may explain both their spontaneous triggering and their long duration. However, as some events are found at shallower depth, the involvement of fluids might not be the only explanation. Other mechanisms such as geometrical complexities might be involved but more evidence are required.

The main outcome of this study is that we found numerous aseismic slip events in a place where none were found during the interseismic period before. As a consequence, aseismic slip events may be found elsewhere in subduction zone contexts where experts did not find any event, pending dedicated noise analysis methods. We provide here one piece of evidence supporting the hypothesis proposed by Jolivet and Frank (2020) which states that slow slip happens everywhere and at all times.

Acknowledgements

The authors thank the Laboratoire International Associé "Montessus de Ballore" (LIA-MB), International Plate Boundary Observatory Chile (IPOC, www.ipoc-network.org), Central Andean Tectonic Observatory Geodetic Array (CAnTO, www.tectonics.caltech.edu/resources/continuous_gps.html), and Instituto Geofísico del Perú (www.igp.gob.pe) for making the raw GNSS data available, as well as all researchers, technicians, and students involved in the installation and maintenance of such networks. J. Jara acknowledges a PhD scholarship granted by Chilean National Science Cooperation (CONICYT) through "Becas Chile" Program and MSCA Postdoctoral Fellowship (MSCA-101066069, project ERASMUS) A. Socquet and J. Jara benefited from a grant from ANR (ANR-17-CE31-0002-01, project AtypicSSE). J. Jara, R. Jolivet and A. Socquet acknowledge the funding of the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement 758210, project Geo4D and grant agreement 865963, project DEEPtrigger). R. Jolivet acknowledges funding from the Institut Universitaire de France. D. Comte acknowledges the funding provided by ANID via the projects PIA/ANID grant ACT-192169, and ANID Project AFB180004, AFB220002. The authors would like to thank J. Pina-Valdes, H. Bhat, H. Sanchez-Reyes, S. Michel, C. Vigny, R. Madariaga, and M. Bouchon for all the constructive discussions about this work. We thank the editors W. Xu, E. Williams, and G. Funning for handling the manuscript, and the reviewers P. Poli and J. Freymueller for their valuable comments, which greatly improved the present manuscript.

Data and code availability

GNSS data are processed using GAMIT software (Herring et al., 2015) (http://geoweb.mit.edu/gg/), while the reference frame is defined using PYACS (Nocquet, 2018) (https://github.com/JMNocquet/pyacs36). GNSS time series used in this work can be found at: https://doi.org/10.5281/zenodo.7898656. The modeling has been performed using the Classic Slip Inversion library (Jolivet et al., 2015b) (CSI, https://github.com/jolivetr/csi) and AlTar (Minson et al., 2013) (https://github.com/AlTarFramework/altar). All plots are made using Matplotlib (Hunter, 2007) and Cartopy (Office, 2010) Python packages.

References

Altamimi, Z., Collilieux, X., and Métivier, L. ITRF2008: An improved solution of the international terrestrial reference frame. *Journal of Geodesy*, 85(8):457–473, 2011. doi: 10.1007/s00190-011-0444-4.

Altamimi, Z., Rebischung, P., Métivier, L., and Collilieux, X. ITRF2014: A new release of the International Terrestrial Reference Frame modeling nonlinear station motions. *Journal of Geophysical Research: Solid Earth*, 121(8):6109–6131, 2016. doi: 10.1002/2016JB013098.

Ambraseys, N. Some characteristic features of the Anatolian fault zone. *Tectonophysics*, 9(2-3):143–165, 3 1970. doi: 10.1016/0040-1951(70)90014-4.

Araki, E., Saffer, D. M., Kopf, A. J., Wallace, L. M., Kimura, T., Machida, Y., Ide, S., and Davis, E. Recurring and triggered slow-slip events near the trench at the Nankai Trough subduction megathrust. *Science*, 356(6343):1157–1160, 2017. doi: 10.1126/science.aan3120.

Armijo, R. and Thiele, R. Active faulting in northern Chile: ramp stacking and lateral decoupling along a subduction plate boundary? *Earth and Planetary Science Letters*, 98(1):40–61, 4 1990. doi: 10.1016/0012-821X(90)90087-E.

Avouac, J.-P. From Geodetic Imaging of Seismic and Aseismic Fault Slip to Dynamic Modeling of the Seismic Cycle. *Annual Review of Earth and Planetary Sciences*, 43(1):233–271, 2015. doi: 10.1146/annurev-earth-060614-105302.

Báez, J. C., Leyton, F., Troncoso, C., del Campo, F., Bevis, M., Vigny, C., Moreno, M., Simons, M., Kendrick, E., Parra, H., and Blume, F. The Chilean GNSS Network: Current Status and Progress toward Early Warning Applications. *Seismological Research Letters*, 89 (4):1546–1554, 7 2018. doi: 10.1785/0220180011.

Bayart, E., Svetlizky, I., and Fineberg, J. Slippery but Tough: The Rapid Fracture of Lubricated Frictional Interfaces. *Physical Review Letters*, 116(19):194301, 5 2016. doi: 10.1103/PhysRevLett.116.194301.

Beck, S. L. and Ruff, L. J. Great earthquakes and subduction along the Peru trench. *Physics of the Earth and Planetary Interiors*, 57 (3-4):199–224, 11 1989. doi: 10.1016/0031-9201(89)90112-X.

Behr, W. M. and Bürgmann, R. What's down there? The structures, materials and environment of deep-seated slow slip and tremor. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2193): 20200218, 3 2021. doi: 10.1098/rsta.2020.0218.

Béjar-Pizarro, M., Carrizo, D., Socquet, A., Armijo, R., Barrientos, S., Bondoux, F., Bonvalot, S., Campos, J., Comte, D., De Chabalier, J. B., and others. Asperities and barriers on the seismogenic zone in North Chile: state-of-the-art after the 2007 Mw 7.7 Tocopilla earthquake inferred by GPS and InSAR data. *Geophysical Journal International*, 183(1):390–406, 2010. doi: 10.1111/j.1365-246X.2010.04748.x.

Béjar-Pizarro, M., Socquet, A., Armijo, R., Carrizo, D., Genrich, J., and Simons, M. Andean structural control on interseismic coupling in the North Chile subduction zone. *Nature Geoscience*, 6 (6):462–467, 6 2013. doi: 10.1038/ngeo1802.

Bevis, M. and Brown, A. Trajectory models and reference frames for crustal motion geodesy. *Journal of Geodesy*, 88(3):283–311, 3 2014. doi: 10.1007/s00190-013-0685-5.

Bock, Y. and Melgar, D. Physical applications of GPS geodesy: A review. *Reports on Progress in Physics*, 79(10):106801, 2016. doi: 10.1088/0034-4885/79/10/106801.

Boehm, J., Werl, B., and Schuh, H. Troposphere mapping functions for GPS and very long baseline interferometry from European Centre for Medium-Range Weather Forecasts operational analysis data. *Journal of Geophysical Research: Solid Earth*, 111(2), 2006. doi: 10.1029/2005JB003629.

Bouchon, M., Marsan, D., Durand, V., Campillo, M., Perfettini, H., Madariaga, R., and Gardonio, B. Potential slab deformation and plunge prior to the Tohoku, Iquique and Maule earthquakes. *Nature Geoscience*, 9(5):380–383, 5 2016. doi: 10.1038/ngeo2701.

Bouchon, M., Marsan, D., Jara, J., Socquet, A., Campillo, M., and Perfettini, H. Suspected Deep Interaction and Triggering Between Giant Earthquakes in the Chilean Subduction Zone. *Geophysical Research Letters*, 45(11):5454–5460, 6 2018. doi: 10.1029/2018GL078350.

Boudin, F., Bernard, P., Meneses, G., Vigny, C., Olcay, M., Tassara, C., Boy, J. P., Aissaoui, E., Métois, M., Satriano, C., Esnoult, M.-F., Nercessian, A., Vallée, M., Vilotte, J.-P., and Brunet, C. Slow slip events precursory to the 2014 Iquique Earthquake, revisited with long-base tilt and GPS records. *Geophysical Journal International*, 228(3):2092–2121, 2021. doi: 10.1093/gji/ggab425.

Brace, W. F. and Byerlee, J. D. Stick-Slip as a Mechanism for Earthquakes. *Science*, 153(3739):990–992, 8 1966. doi: 10.1126/science.153.3739.990.

Bürgmann, R. The geophysics, geology and mechanics of slow fault slip. *Earth and Planetary Science Letters*, 495:112–134, 2018. doi: 10.1016/j.epsl.2018.04.062.

Bürgmann, R., Kogan, M. G., Levin, V. E., Scholz, C. H., King, R. W., and Steblov, G. M. Rapid aseismic moment release following the 5 December, 1997 Kronotsky, Kamchatka, earthquake. *Geophysical Research Letters*, 28(7):1331–1334, 4 2001. doi: 10.1029/2000GL012350.

Bürgmann, R., Kogan, M. G., Steblov, G. M., Hilley, G., Levin, V. E., and Apel, E. Interseismic coupling and asperity distribution along the Kamchatka subduction zone. *Journal of Geophysical Research: Solid Earth*, 110(7):1–17, 2005. doi: 10.1029/2005JB003648.

Cheloni, D., D'Agostino, N., Selvaggi, G., Avallone, A., Fornaro, G., Giuliani, R., Reale, D., Sansosti, E., and Tizzani, P. Aseismic transient during the 2010–2014 seismic swarm: evidence for longer recurrence of M > 6.5 earthquakes in the Pollino gap (Southern Italy)? *Scientific Reports*, 7(1):576, 4 2017. doi: 10.1038/s41598-017-00649-z.

Chlieh, M., De Chabalier, J. B., Ruegg, J. C., Armijo, R., Dmowska, R., Campos, J., and Feigl, K. L. Crustal deformation and fault slip during the seismic cycle in the North Chile subduction zone, from GPS and InSAR observations. *Geophysical Journal International*, 158(2):695–711, 2004. doi: 10.1111/j.1365-246X.2004.02326.x.

Chlieh, M., Perfettini, H., Tavera, H., Avouac, J. P., Remy, D., Nocquet, J. M., Rolandone, F., Bondoux, F., Gabalda, G., and Bonvalot, S. Interseismic coupling and seismic potential along the Central Andes subduction zone. *Journal of Geophysical Research: Solid Earth*, 116(12), 2011. doi: 10.1029/2010JB008166.

Comte, D. and Pardo, M. Reappraisal of great historical earthquakes in the northern Chile and southern Peru seismic gaps. *Natural Hazards*, 4(1):23–44, 1991. doi: 10.1007/BF00126557.

Comte, D., Carrizo, D., Roecker, S., Ortega-Culaciati, F., and Peyrat, S. Three-dimensional elastic wave speeds in the northern Chile subduction zone: Variations in hydration in the supraslab mantle. *Geophysical Journal International*, 207(2):1080–1105, 2016. doi: 10.1093/gji/ggw318.

Contreras-Reyes, E., Díaz, D., Bello-González, J. P., Slezak, K., Potin, B., Comte, D., Maksymowicz, A., Ruiz, J. A., Osses, A., and Ruiz, S. Subduction zone fluids and arc magmas conducted by lithospheric deformed regions beneath the central Andes. *Scientific Reports*, 11(1):1–12, 2021. doi: 10.1038/s41598-021-02430-9.

Dal Zilio, L., Jolivet, R., and van Dinther, Y. Segmentation of the Main Himalayan Thrust Illuminated by Bayesian Inference of Interseismic Coupling. *Geophysical Research Letters*, 47(4):1–10, 2020a. doi: 10.1029/2019GL086424.

Dal Zilio, L., Lapusta, N., and Avouac, J. P. Unraveling Scaling Properties of Slow-Slip Events. *Geophysical Research Letters*, 47(10), 2020b. doi: 10.1029/2020GL087477.

Delouis, B., Monfret, T., Dorbath, L., Pardo, M., Rivera, L., Comte, D., Haessler, H., Caminade, J. P., Ponce, L., Kausel, E., and Cisternas, A. The Mw= 8.0 antofagasta (northern Chile) earthquake of 30 July 1995: A precursor to the end of the large 1877 gap. *Bulletin of the Seismological Society of America*, 87(2):427–445, 1997.

Dorbath, L., Cisternas, A., and Dorbath, C. Assessment of the size of large and great historical earthquakes in Peru. *Bulletin of the Seismological Society of America*, 80(3):551–576, 1990.

Dragert, H., Wang, K., and James, T. S. A Silent Slip Event on the Deeper Cascadia Subduction Interface. *Science*, 292(5521): 1525–1528, 5 2001. doi: 10.1126/science.1060152.

Ducellier, A., Creager, K. C., and Schmidt, D. A. Detection of Slow Slip Events Using Wavelet Analysis of GNSS Recordings. *Bulletin of the Seismological Society of America*, 112(5):2408–2424, 2022. doi: 10.1785/0120210289.

Duputel, Z., Agram, P. S., Simons, M., Minson, S. E., and Beck, J. L. Accounting for prediction uncertainty when inferring subsurface fault slip. *Geophysical Journal International*, 197(1): 464–482, 4 2014. doi: 10.1093/gji/ggt517.

Duputel, Z., Jiang, J., Jolivet, R., Simons, M., Rivera, L., Ampuero, J. P., Riel, B., Owen, S. E., Moore, A. W., Samsonov, S. V., Ortega Culaciati, F., and Minson, S. E. The Iquique earthquake sequence of April 2014: Bayesian modeling accounting for prediction uncertainty. *Geophysical Research Letters*, 42(19): 7949–7957, 2015. doi: 10.1002/2015GL065402.

Dziewonski, A. M., Chou, T.-A., and Woodhouse, J. H. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *Journal of Geophysical Research: Solid Earth*, 86(B4):2825–2852, 4 1981. doi: 10.1029/JB086iB04p02825.

Ekström, G., Nettles, M., and Dziewoński, A. The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors*, 200-201:1–9, 6 2012. doi: 10.1016/j.pepi.2012.04.002.

Essing, D. and Poli, P. Spatiotemporal Evolution of the Seismicity in the Alto Tiberina Fault System Revealed by a High-Resolution Template Matching Catalog. *Journal of Geophysical Research: Solid Earth*, 127(10), 10 2022. doi: 10.1029/2022JB024845.

Frank, W. B. Slow slip hidden in the noise: The intermittence of tectonic release. *Geophysical Research Letters*, 43(19):125–10, 2016. doi: 10.1002/2016GL069537.

Frank, W. B. and Brodsky, E. E. Daily measurement of slow slip from low-frequency earthquakes is consistent with ordinary earthquake scaling. *Science Advances*, 5(10):eaaw9386, 10 2019. doi:

10.1126/sciadv.aaw9386.

Gardonio, B., Marsan, D., Socquet, A., Bouchon, M., Jara, J., Sun, Q., Cotte, N., and Campillo, M. Revisiting Slow Slip Events Occurrence in Boso Peninsula, Japan, Combining GPS Data and Repeating Earthquakes Analysis. *Journal of Geophysical Research: Solid Earth*, 123(2):1502–1515, 2 2018. doi: 10.1002/2017JB014469.

Gomberg, J., Wech, A., Creager, K., Obara, K., and Agnew, D. Reconsidering earthquake scaling. *Geophysical Research Letters*, 43(12):6243–6251, 6 2016. doi: 10.1002/2016GL069967.

Graham, S., DeMets, C., Cabral-Cano, E., Kostoglodov, V., Rousset, B., Walpersdorf, A., Cotte, N., Lasserre, C., McCaffrey, R., and Salazar-Tlaczani, L. Slow Slip History for the MEXICO Subduction Zone: 2005 Through 2011. *Pure and Applied Geophysics*, 173(10-11):3445–3465, 2016. doi: 10.1007/s00024-015-1211-x.

Gualandi, A., Nichele, C., Serpelloni, E., Chiaraluce, L., Anderlini, L., Latorre, D., Belardinelli, M. E., and Avouac, J.-P. Aseismic deformation associated with an earthquake swarm in the northern Apennines (Italy). *Geophysical Research Letters*, 44(15): 7706–7714, 8 2017. doi: 10.1002/2017GL073687.

Harris, R. A. Large earthquakes and creeping faults. *Reviews of Geophysics*, 55(1):169–198, 2017. doi: 10.1002/2016RG000539.

Hartzell, S. and Langer, C. Importance of model parameterization in finite fault inversions: application to the 1974 MW 8.0 Peru earthquake. *Journal of Geophysical Research*, 98(B12), 1993. doi: 10.1029/93jb02453.

Hayes, G. P., Moore, G. L., Portner, D. E., Hearne, M., Flamme, H., Furtney, M., and Smoczyk, G. M. Slab2, a comprehensive subduction zone geometry model. *Science*, 362(6410):58–61, 10 2018. doi: 10.1126/science.aat4723.

Heki, K., Miyazaki, S., and Tsuji, H. Silent fault slip following an interplate thrust earthquake at the Japan Trench. *Nature*, 386 (6625):595–598, 4 1997. doi: 10.1038/386595a0.

Herring, T. A., King, R., Floyd, M., and McClusky, S. C. GAMIT Reference Manual. GPS Analysis at MIT GLOBK, Release 10.6, 2015.

Hetland, E. A. and Simons, M. Post-seismic and interseismic fault creep II: Transient creep and interseismic stress shadows on megathrusts. *Geophysical Journal International*, 181(1):99–112, 2010. doi: 10.1111/j.1365-246X.2009.04482.x.

Hino, R., Inazu, D., Ohta, Y., Ito, Y., Suzuki, S., Iinuma, T., Osada, Y., Kido, M., Fujimoto, H., and Kaneda, Y. Was the 2011 Tohoku-Oki earthquake preceded by aseismic preslip? Examination of seafloor vertical deformation data near the epicenter. *Marine Geophysical Research*, 35(3):181–190, 2014. doi: 10.1007/s11001-013-9208-2.

Hirose, H., Hirahara, K., Kimata, F., Fujii, N., and Miyazaki, S. A slow thrust slip event following the two 1996 Hyuganada Earthquakes beneath the Bungo Channel, southwest Japan. *Geophysical Research Letters*, 26(21):3237–3240, 11 1999. doi: 10.1029/1999GL010999.

Hoffmann, F., Metzger, S., Moreno, M., Deng, Z., Sippl, C., Ortega-Culaciati, F., and Oncken, O. Characterizing Afterslip and Ground Displacement Rate Increase Following the 2014 Iquique-Pisagua Mw8.1 Earthquake, Northern Chile. *Journal of Geophysical Research: Solid Earth*, 123(5):4171–4192, 5 2018. doi: 10.1002/2017JB014970.

Hsu, Y.-J., Bechor, N., Segall, P., Yu, S.-B., Kuo, L.-C., and Ma, K.-F. Rapid afterslip following the 1999 Chi-Chi, Taiwan Earthquake. *Geophysical Research Letters*, 29(16):1–4, 8 2002. doi: 10.1029/2002GL014967.

Hsu, Y. J., Simons, M., Avouac, J. P., Galeteka, J., Sieh, K., Chlieh, M., Natawidjaja, D., Prawirodirdjo, L., and Bock, Y. Frictional afterslip following the 2005 Nias-Simeulue earthquake, Sumatra. *Science*, 312(5782):1921–1926, 2006. doi: 10.1126/science.1126960.

Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, 9(3):90–95, 2007. doi: 10.1109/MCSE.2007.55.

Husen, S., Kissling, E., Flueh, E., and Asch, G. Accurate hypocentre determination in the seismogenic zone of the subducting Nazca Plate in northern Chile using a combined on-/offshore network. *Geophysical Journal International*, 138(3):687–701, 1999. doi: 10.1046/j.1365-246X.1999.00893.x.

Ide, S. and Beroza, G. C. Slow earthquake scaling reconsidered as a boundary between distinct modes of rupture propagation. *Proceedings of the National Academy of Sciences*, 120(32):2017, 8 2023. doi: 10.1073/pnas.2222102120.

Ide, S., Beroza, G. C., Shelly, D. R., and Uchide, T. A scaling law for slow earthquakes. *Nature*, 447(7140):76–79, 2007. doi: 10.1038/nature05780.

International Seismological Centre. On-line Bulletin, 2016.

Itoh, Y., Aoki, Y., and Fukuda, J. Imaging evolution of Cascadia slow-slip event using high-rate GPS. *Scientific Reports*, 12(1): 1–12, 2022. doi: 10.1038/s41598-022-10957-8.

Jara, J., Socquet, A., Marsan, D., and Bouchon, M. Long-Term Interactions Between Intermediate Depth and Shallow Seismicity in North Chile Subduction Zone. *Geophysical Research Letters*, 44(18):9283–9292, 9 2017. doi: 10.1002/2017GL075029.

Jara, J., Sánchez-Reyes, H., Socquet, A., Cotton, F., Virieux, J., Maksymowicz, A., Díaz-Mojica, J., Walpersdorf, A., Ruiz, J., Cotte, N., and Norabuena, E. Kinematic study of Iquique 2014 M w 8.1 earthquake: Understanding the segmentation of the seismogenic zone. *Earth and Planetary Science Letters*, 503: 131–143, 2018. doi: 10.1016/j.epsl.2018.09.025.

Jolivet, R. and Frank, W. B. The Transient and Intermittent Nature of Slow Slip. *AGU Advances*, 1(1), 2020. doi: 10.1029/2019av000126.

Jolivet, R. and Simons, M. A Multipixel Time Series Analysis Method Accounting for Ground Motion, Atmospheric Noise, and Orbital Errors. *Geophysical Research Letters*, 45(4):1814–1824, 2018. doi: 10.1002/2017GL076533.

Jolivet, R., Candela, T., Lasserre, C., Renard, F., Klinger, Y., and Doin, M. The Burst-Like Behavior of Aseismic Slip on a Rough Fault: The Creeping Section of the Haiyuan Fault, China. *Bulletin of the Seismological Society of America*, 105(1):480–488, 2 2015a. doi: 10.1785/0120140237.

Jolivet, R., Simons, M., Agram, P. S., Duputel, Z., and Shen, Z. K. Aseismic slip and seismogenic coupling along the central San Andreas Fault. *Geophysical Research Letters*, 42(2):297–306, 2015b. doi: 10.1002/2014GL062222.

Jolivet, R., Simons, M., Duputel, Z., Olive, J. A., Bhat, H. S., and Bletery, Q. Interseismic Loading of Subduction Megathrust Drives Long-Term Uplift in Northern Chile. *Geophysical Research Letters*, 47(8):1–11, 2020. doi: 10.1029/2019GL085377.

Kanamori, H. The Nature of Seismicity Patterns Before Large Earthquakes. In *Earthquake Prediction*, pages 1–19. Wiley Online Library, 3 1981. doi: 10.1029/ME004p0001.

Kato, A. and Ben-Zion, Y. The generation of large earthquakes. *Nature Reviews Earth & Environment*, 2(1):26–39, 1 2021. doi: 10.1038/s43017-020-00108-w.

Kato, A. and Nakagawa, S. Multiple slow-slip events during a foreshock sequence of the 2014 Iquique, Chile Mw 8.1 earthquake. *Geophysical Research Letters*, 41(15):5420–5427, 2014. doi: 10.1002/2014GL061138.

Kato, A., Obara, K., Igarashi, T., Tsuruoka, H., Nakagawa, S., and Hirata, N. Propagation of slow slip leading up to the 2011 Mw9.0

Tohoku-Oki earthquake. *Science*, 335(6069):705–708, 2 2012. doi: 10.1126/science.1215141.

Kausel, E. Los terremotos de agosto de 1868 y mayo de 1877 que afectaron el sur del Perú y norte de Chile. *Boletín de la Academia Chilena de Ciencias*, 3(1):8–13, 1986.

Khoshmanesh, M. and Shirzaei, M. Episodic creep events on the San Andreas Fault caused by pore pressure variations. *Nature Geoscience*, 11(8):610–614, 2018. doi: 10.1038/s41561-018-0160-2.

Klein, E., Vigny, C., Nocquet, J. M., and Boulze, H. A 20 year-long GNSS solution across South-America with focus in Chile. *BSGF - Earth Sciences Bulletin*, 193, 2022. doi: 10.1051/bsgf/2022005.

Klotz, J., Deng, Z., Moreno, M., Asch, G., Bartsch, M., and Ramatschi, M. IPOC cGPS - Continuous Mode GPS data in the IPOC Region, Northern Chile. Technical report, GFZ Data Services, 2017.

Lay, T. The surge of great earthquakes from 2004 to 2014. *Earth and Planetary Science Letters*, 409(October 2016):133–146, 2015. doi: 10.1016/j.epsl.2014.10.047.

Li, Y., Nocquet, J. M., Shan, X., and Song, X. Geodetic Observations of Shallow Creep on the Laohushan-Haiyuan Fault, Northeastern Tibet. *Journal of Geophysical Research: Solid Earth*, 126(6): 1–18, 2021. doi: 10.1029/2020JB021576.

Lindsey, E. O., Mallick, R., Hubbard, J. A., Bradley, K. E., Almeida, R. V., Moore, J. D. P., Bürgmann, R., and Hill, E. M. Slip rate deficit and earthquake potential on shallow megathrusts. *Nature Geoscience*, 14(5):321–326, 5 2021. doi: 10.1038/s41561-021-00736-x.

Liu, Y. and Rice, J. R. Aseismic slip transients emerge spontaneously in three-dimensional rate and state modeling of subduction earthquake sequences. *Journal of Geophysical Research: Solid Earth*, 110(8):1–14, 2005. doi: 10.1029/2004JB003424.

Liu, Y. and Rice, J. R. Spontaneous and triggered aseismic deformation transients in a subduction fault model. *Journal of Geophysical Research: Solid Earth*, 112(9), 2007. doi: 10.1029/2007JB004930.

Louderback, G. Faults and Eartquakes. *Bulletin of Seismological Society of America*, 32(4):305–330, 1942.

Loveless, J. P. and Meade, B. J. Geodetic imaging of plate motions, slip rates, and partitioning of deformation in Japan. *Journal of Geophysical Research*, 115(B2):B02410, 2010. doi: 10.1029/2008JB006248.

Lovery, B., Chlieh, M., Norabuena, E., Villegas-Lanza, J. C., Radiguet, M., Cotte, N., Tsapong-Tsague, A., Quiroz, W., Sierra Farfán, C., Simons, M., Nocquet, J. M., Tavera, H., and Socquet, A. Heterogeneous Locking and Earthquake Potential on the South Peru Megathrust From Dense GNSS Network. *Journal of Geophysical Research: Solid Earth*, 129(2), 2 2024. doi: 10.1029/2023JB027114.

Marsan, D., Reverso, T., Helmstetter, A., and Enescu, B. Slow slip and aseismic deformation episodes associated with the subducting Pacific plate offshore Japan, revealed by changes in seismicity. *Journal of Geophysical Research E: Planets*, 118(9): 4900–4909, 2013. doi: 10.1002/jgrb.50323.

Marsan, D., Bouchon, M., Gardonio, B., Perfettini, H., Socquet, A., and Enescu, B. Change in seismicity along the Japan trench, 1990-2011, and its relationship with seismic coupling. *Journal of Geophysical Research: Solid Earth*, 122(6):4645–4659, 6 2017. doi: 10.1002/2016JB013715.

Materna, K., Bartlow, N., Wech, A., Williams, C., and Bürgmann, R. Dynamically Triggered Changes of Plate Interface Coupling in Southern Cascadia. *Geophysical Research Letters*, 46(22): 12890–12899, 2019. doi: 10.1029/2019GL084395.

Mazzotti, S. S., Le Pichon, X., Henry, P., and Miyazaki, S.-I. Full interseismic locking of the Nankai and Japan-west Kurile subduction zones: An analysis of uniform elastic strain accumulation in Japan constrained by permanent GPS. *Journal of Geophysical Research: Solid Earth*, 105(B6):13159–13177, 2000. doi: 10.1029/2000jb900060.

McLaskey, G. C. Earthquake Initiation From Laboratory Observations and Implications for Foreshocks. *Journal of Geophysical Research: Solid Earth*, 124(12):12882–12904, 2019. doi: 10.1029/2019JB018363.

Melbourne, T. I. Precursory transient slip during the 2001 M w = 8.4 Peru earthquake sequence from continuous GPS. *Geophysical Research Letters*, 29(21):2032, 2002. doi: 10.1029/2002GL015533.

Melnick, D., Moreno, M., Quinteros, J., Baez, J. C., Deng, Z., Li, S., and Oncken, O. The super-interseismic phase of the megathrust earthquake cycle in Chile. *Geophysical Research Letters*, 44(2): 784–791, 2017. doi: 10.1002/2016gl071845.

Meng, L., Huang, H., Bürgmann, R., Ampuero, J. P., and Strader, A. Dual megathrust slip behaviors of the 2014 Iquique earthquake sequence. *Earth and Planetary Science Letters*, 411:177–187, 2015. doi: 10.1016/j.epsl.2014.11.041.

Métois, M., Vigny, C., and Socquet, A. Interseismic Coupling, Megathrust Earthquakes and Seismic Swarms Along the Chilean Subduction Zone (38°–18°S). *Pure and Applied Geophysics*, 173 (5):1431–1449, 2016. doi: 10.1007/s00024-016-1280-5.

Michel, S., Gualandi, A., and Avouac, J. P. Interseismic Coupling and Slow Slip Events on the Cascadia Megathrust. *Pure and Applied Geophysics*, 176(9):3867–3891, 2019a. doi: 10.1007/s00024-018-1991-x.

Michel, S., Gualandi, A., and Avouac, J.-P. Similar scaling laws for earthquakes and Cascadia slow-slip events. *Nature*, 574(7779): 522–526, 10 2019b. doi: 10.1038/s41586-019-1673-6.

Michel, S., Jolivet, R., Lengliné, O., Gualandi, A., Larochelle, S., and Gardonio, B. Searching for Transient Slow Slips Along the San Andreas Fault Near Parkfield Using Independent Component Analysis. *Journal of Geophysical Research: Solid Earth*, 127 (6):1–19, 2022. doi: 10.1029/2021JB023201.

Minson, S. E., Simons, M., and Beck, J. L. Bayesian inversion for finite fault earthquake source models I-theory and algorithm. *Geophysical Journal International*, 194(3):1701–1726, 2013. doi: 10.1093/gji/ggt180.

Müller, R. D., Sdrolias, M., Gaina, C., and Roest, W. R. Age, spreading rates, and spreading asymmetry of the world's ocean crust. *Geochemistry, Geophysics, Geosystems*, 9(4), 4 2008. doi: 10.1029/2007GC001743.

Nishikawa, T., Matsuzawa, T., Ohta, K., Uchida, N., Nishimura, T., and Ide, S. The slow earthquake spectrum in the Japan Trench illuminated by the S-net seafloor observatories. *Science*, 365 (6455):808–813, 2019. doi: 10.1126/science.aax5618.

Nishimura, T. Short-term slow slip events along the Ryukyu Trench, southwestern Japan, observed by continuous GNSS. *Progress in Earth and Planetary Science*, 1(1):1–13, 2014. doi: 10.1186/s40645-014-0022-5.

Nishimura, T., Matsuzawa, T., and Obara, K. Detection of short-term slow slip events along the Nankai Trough, southwest Japan, using GNSS data. *Journal of Geophysical Research: Solid Earth*, 118 (6):3112–3125, 2013. doi: 10.1002/jgrb.50222.

Nocquet, J. M. PYACS: A set of Python tools for GPS analysis and tectonic modelling. In *PYACS: A set of Python tools for GPS analysis and tectonic modelling*. 19th General Assembly of Wegener, 2018.

Nocquet, J. M., Villegas-Lanza, J. C., Chlieh, M., Mothes, P. A., Rolandone, F., Jarrin, P., Cisneros, D., Alvarado, A., Audin, L., Bondoux, F., Martin, X., Font, Y., Régnier, M., Vallée, M., Tran, T., Beauval, C., Maguiña Mendoza, J. M., Martinez, W., Tavera, H., and Yepes, H. Motion of continental slivers and creeping subduction in the northern Andes. *Nature Geoscience*, 7(4):287–291, 2014. doi: 10.1038/ngeo2099.

Obara, K. and Kato, A. Connecting slow earthquakes to huge earthquakes. *Science (New York, N.Y.)*, 353(6296):253–257, 2016. doi: 10.1126/science.aaf1512.

Office, M. Cartopy: a cartographic python library with a Matplotlib interface, 2010.

Peacock, S. M. Are the lower planes of double seismic zones caused by serpentine dehydration in subducting oceanic mantle? *Geology*, 29(4):299–302, 2001. doi: 10.1130/0091-7613(2001)029<0299:ATLPOD>2.0.CO;2.

Peng, Z. and Gomberg, J. An integrated perspective of the continuum between earthquakes and slow-slip phenomena. *Nature Geoscience*, 3(9):599–607, 2010. doi: 10.1038/ngeo940.

Perfettini, H. and Ampuero, J. P. Dynamics of a velocity strengthening fault region: Implications for slow earthquakes and postseismic slip. *Journal of Geophysical Research: Solid Earth*, 113 (9), 2008. doi: 10.1029/2007JB005398.

Perfettini, H., Avouac, J. P., Tavera, H., Kositsky, A., Nocquet, J. M., Bondoux, F., Chlieh, M., Sladen, A., Audin, L., Farber, D. L., and Soler, P. Seismic and aseismic slip on the Central Peru megathrust. *Nature*, 465(7294):78–81, 2010. doi: 10.1038/nature09062.

Peyrat, S. and Favreau, P. Kinematic and spontaneous rupture models of the 2005 Tarapacá intermediate depth earthquake. *Geophysical Journal International*, 181(1):369–381, 2010. doi: 10.1111/j.1365-246X.2009.04493.x.

Peyrat, S., Campos, J., de Chabalier, J. B., Perez, A., Bonvalot, S., Bouin, M. P., Legrand, D., Nercessian, A., Charade, O., Patau, G., Clévédæ, E., Kausel, E., Bernard, P., and Vilotte, J. P. Tarapacá intermediate-depth earthquake (Mw 7.7, 2005, northern Chile): A slab-pull event with horizontal fault plane constrained from seismologic and geodetic observations. *Geophysical Research Letters*, 33(22):1–6, 2006. doi: 10.1029/2006GL027710.

Poli, P., Jeria, A. M., and Ruiz, S. The M w 8.3 Illapel earthquake (Chile): Preseismic and postseismic activity associated with hydrated slab structures. *Geology*, 45(3):247–250, 3 2017. doi: 10.1130/G38522.1.

Pritchard, M. E. and Simons, M. An aseismic slip pulse in northern Chile and along-strike variations in seismogenic behavior. *Journal of Geophysical Research: Solid Earth*, 111(8), 2006. doi: 10.1029/2006JB004258.

Pritchard, M. E., Norabuena, E. O., Ji, C., Boroschek, R., Comte, D., Simons, M., Dixon, T. H., and Rosen, P. A. Geodetic, teleseismic, and strong motion constraints on slip from recent southern Peru subduction zone earthquakes. *Journal of Geophysical Research: Solid Earth*, 112(3), 2007. doi: 10.1029/2006JB004294.

Radiguet, M., Cotton, F., Vergnolle, M., Campillo, M., Walpersdorf, A., Cotte, N., and Kostoglodov, V. Slow slip events and strain accumulation in the Guerrero gap, Mexico. *Journal of Geophysical Research: Solid Earth*, 117(4), 2012. doi: 10.1029/2011JB008801.

Radiguet, M., Perfettini, H., Cotte, N., Gualandi, A., Valette, B., Kostoglodov, V., Lhomme, T., Walpersdorf, A., Cabral Cano, E., and Campillo, M. Triggering of the 2014 Mw7.3 Papanoa earthquake by a slow slip event in Guerrero, Mexico. *Nature Geoscience*, 9 (11):829–833, 2016. doi: 10.1038/ngeo2817.

Reid, H. F. The Mechanism of the Earthquake. The California Earthquake of April 18, 1906: Rep. of the State Investigation Commiss. Vol. 2. P. 1. Technical report, Carnigie Institution of Washington, Washington, D. C., 1910.

Remy, D., Perfettini, H., Cotte, N., Avouac, J. P., Chlieh, M., Bondoux, F., Sladen, A., Tavera, H., and Socquet, A. Postseismic relocking of the subduction megathrust following the 2007 Pisco, Peru, earthquake. *Journal of Geophysical Research: Solid Earth*, 121(5):3978–3995, 5 2016. doi: 10.1002/2015JB012417.

Reverso, T., Marsan, D., Helmstetter, A., and Enescu, B. Background seismicity in Boso Peninsula, Japan: Long-term acceleration, and relationship with slow slip events. *Geophysical Research Letters*, 43(11):5671–5679, 2016. doi: 10.1002/2016GL068524.

Romanet, P., Bhat, H. S., Jolivet, R., and Madariaga, R. Fast and Slow Slip Events Emerge Due to Fault Geometrical Complexity. *Geophysical Research Letters*, 45(10):4809–4819, 2018. doi: 10.1029/2018GL077579.

Rousset, B., Campillo, M., Lasserre, C., Frank, W. B., Cotte, N., Walpersdorf, A., Socquet, A., and Kostoglodov, V. A geodetic matched filter search for slow slip with application to the Mexico subduction zone. *Journal of Geophysical Research: Solid Earth*, 122(12):498–10, 12 2017. doi: 10.1002/2017JB014448.

Rousset, B., Bürgmann, R., and Campillo, M. Slow slip events in the roots of the San Andreas fault. *Science Advances*, 5(2):eaav3274, 2 2019. doi: 10.1126/sciadv.aav3274.

Ruegg, J. C., Olcay, M., and Lazo, D. Co-, Post- and Pre(?)-seismic Displacements Associated with the Mw 8.4 Southern Peru Earthquake of 23 June 2001 from Continuous GPS Measurements. *Seismological Research Letters*, 72(6):673–678, 11 2001. doi: 10.1785/gssrl.72.6.673.

Ruiz, S. and Madariaga, R. Historical and recent large megathrust earthquakes in Chile. *Tectonophysics*, 733(September 2017): 37–56, 2018. doi: 10.1016/j.tecto.2018.01.015.

Ruiz, S., Metois, M., Fuenzalida, A., Ruiz, J., Leyton, F., Grandin, R., Vigny, C., Madariaga, R., and Campos, J. Intense foreshocks and a slow slip event preceded the 2014 Iquique Mw8.1 earthquake. *Science*, 345(6201):1165–1169, 2014. doi: 10.1126/science.1256074.

Ruiz, S., Klein, E., del Campo, F., Rivera, E., Poli, P., Metois, M., Christophe, V., Baez, J. C., Vargas, G., Leyton, F., Madariaga, R., and Fleitout, L. The Seismic Sequence of the 16 September 2015 M w 8.3 Illapel, Chile, Earthquake. *Seismological Research Letters*, 87(4):789–799, 7 2016. doi: 10.1785/0220150281.

Ruiz, S., Moreno, M., Melnick, D., del Campo, F., Poli, P., Baez, J. C., Leyton, F., and Madariaga, R. Reawakening of large earthquakes in south central Chile: The 2016 Mw7.6 Chiloé event. *Geophysical Research Letters*, 44(13):6633–6640, 7 2017. doi: 10.1002/2017GL074133.

Rüpke, L. H., Morgan, J. P., Hort, M., and Connolly, J. A. Serpentine and the subduction zone water cycle. *Earth and Planetary Science Letters*, 223(1-2):17–34, 2004. doi: 10.1016/j.epsl.2004.04.018.

Savage, J. C. A dislocation model of strain accumulation and release at a subduction zone. *Journal of Geophysical Research: Solid Earth*, 88(B6):4984–4996, 6 1983. doi: 10.1029/JB088iB06p04984.

Schurr, B., Asch, G., Hainzl, S., Bedford, J., Hoechner, A., Palo, M., Wang, R., Moreno, M., Bartsch, M., Zhang, Y., Oncken, O., Tilmann, F., Dahm, T., Victor, P., Barrientos, S., and Vilotte, J.-P. Gradual unlocking of plate boundary controlled initiation of the 2014 Iquique earthquake. *Nature*, 512(7514):299–302, 8 2014. doi: 10.1038/nature13681.

Schurr, B., Moreno, M., Tréhu, A. M., Bedford, J., Kummerow, J., Li, S., and Oncken, O. Forming a Mogi Doughnut in the Years Prior to and Immediately Before the 2014 M8.1 Iquique, Northern Chile, Earthquake. *Geophysical Research Letters*, 47(16), 2020. doi: 10.1029/2020GL088351.

Shrivastava, M. N., González, G., Moreno, M., Soto, H., Schurr, B., Salazar, P., and Báez, J. C. Earthquake segmentation in northern Chile correlates with curved plate geometry. *Scientific Reports*, 9(1):4403, 12 2019. doi: 10.1038/s41598-019-40282-6.

Simons, M., Galetzka, J. E., Genrich, J. F., Ortega, F., Comte, D., Glass, B., Gonzalez, G., and Norabuena, E. Central Andean Tectonic Observatory Geodetic Array - GPS/GNSS Observations. Technical report, Caltech, 2010.

Sippl, C., Schurr, B., Asch, G., and Kummerow, J. Seismicity Structure of the Northern Chile Forearc From >100,000 Double-Difference Relocated Hypocenters\. *Journal of Geophysical Research: Solid Earth*, 123(5):4063–4087, 2018. doi: 10.1002/2017JB015384.

Sippl, C., Schurr, B., Münchmeyer, J., Barrientos, S., and Oncken, O. The Northern Chile forearc constrained by 15 years of permanent seismic monitoring. *Journal of South American Earth Sciences*, 126(December 2022):104326, 6 2023. doi: 10.1016/j.jsames.2023.104326.

Sladen, A., Tavera, H., Simons, M., Avouac, J. P., Konca, A. O., Perfettini, H., Audin, L., Fielding, E. J., Ortega, F., and Cavagnoud, R. Source model of the 2007 Mw8.0 Pisco, Peru earthquake: Implications for seismogenic behavior of subduction megathrusts. *Journal of Geophysical Research: Solid Earth*, 115(2), 2010. doi: 10.1029/2009JB006429.

Socquet, A., Valdes, J. P., Jara, J., Cotton, F., Walpersdorf, A., Cotte, N., Specht, S., Ortega-Culaciati, F., Carrizo, D., and Norabuena, E. An 8 month slow slip event triggers progressive nucleation of the 2014 Chile megathrust. *Geophysical Research Letters*, 44(9): 4046–4053, 5 2017. doi: 10.1002/2017GL073023.

Steinbrugge, K. V., Zacher, E. G., Tocher, D., Whitten, C. A., and Claire, C. N. Creep on the San Andreas fault. *Bulletin of the Seismological Society of America*, 50(3):389–415, 7 1960.

Supino, M., Poiata, N., Festa, G., Vilotte, J. P., Satriano, C., and Obara, K. Self-similarity of low-frequency earthquakes. *Scientific Reports*, 10(1):6523, 12 2020. doi: 10.1038/s41598-020-63584-6.

Takagi, R., Uchida, N., and Obara, K. Along-Strike Variation and Migration of Long-Term Slow Slip Events in the Western Nankai Subduction Zone, Japan. *Journal of Geophysical Research: Solid Earth*, 124(4):3853–3880, 2019. doi: 10.1029/2018JB016738.

Tarantola, A. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, 2005. doi: 10.1137/1.9780898717921.

Tassara, A. and Echaurren, A. Anatomy of the Andean subduction zone: three-dimensional density model upgraded and compared against global-scale models. *Geophysical Journal International*, 189(1):161–168, 4 2012. doi: 10.1111/j.1365-246X.2012.05397.x.

Teunissen, P. J. and Montenbruck, O., editors. *Springer Handbook of Global Navigation Satellite Systems*. Springer International Publishing, Cham, 2017. doi: 10.1007/978-3-319-42928-1.

Tissandier, R., Nocquet, J., Klein, E., Vigny, C., Ojeda, J., and Ruiz, S. Afterslip of the M w 8.3 2015 Illapel Earthquake Imaged Through a Time-Dependent Inversion of Continuous and Survey GNSS Data. *Journal of Geophysical Research: Solid Earth*, 128(2):1–21, 2 2023. doi: 10.1029/2022JB024778.

Twardzik, C., Duputel, Z., Jolivet, R., Klein, E., and Rebischung, P. Bayesian inference on the initiation phase of the 2014 Iquique, Chile, earthquake. *Earth and Planetary Science Letters*, 600: 117835, 2022. doi: 10.1016/j.epsl.2022.117835.

Uchida, N., Takagi, R., Asano, Y., and Obara, K. Migration of shallow and deep slow earthquakes toward the locked segment of the Nankai megathrust. *Earth and Planetary Science Letters*, 531: 115986, 2020. doi: 10.1016/j.epsl.2019.115986.

van Rijsingen, E. M., Calais, E., Jolivet, R., de Chabalier, J., Jara, J., Symithe, S., Robertson, R., and Ryan, G. A. Inferring Interseismic Coupling Along the Lesser Antilles Arc: A Bayesian Approach. *Journal of Geophysical Research: Solid Earth*, 126(2): 1–21, 2 2021. doi: 10.1029/2020JB020677.

Vigny, C. and Klein, E. The 1877 megathrust earthquake of North Chile two times smaller than thought? A review of ancient articles. *Journal of South American Earth Sciences*, 117:103878, 8 2022. doi: 10.1016/j.jsames.2022.103878.

Villegas-Lanza, J. C., Chlieh, M., Cavalié, O., Tavera, H., Baby, P., Chire-Chira, J., and Nocquet, J.-M. Active tectonics of Peru: Heterogeneous interseismic coupling along the Nazca megathrust, rigid motion of the Peruvian Sliver, and Subandean shortening accommodation. *Journal of Geophysical Research: Solid Earth*, 121(10):7371–7394, 10 2016. doi: 10.1002/2016JB013080.

Voss, N., Dixon, T. H., Liu, Z., Malservisi, R., Protti, M., and Schwartz, S. Do slow slip events trigger large and great megathrust earthquakes? *Science advances*, 4(10):eaat8472, 10 2018. doi: 10.1126/sciadv.aat8472.

Wallace, L. M. Slow Slip Events in New Zealand. *Annual Review of Earth and Planetary Sciences*, 48(1):175–203, 5 2020. doi: 10.1146/annurev-earth-071719-055104.

Wang, H., Huismans, R. S., and Rondenay, S. Water Migration in the Subduction Mantle Wedge: A Two-Phase Flow Approach. *Journal of Geophysical Research: Solid Earth*, 124(8):9208–9225, 2019. doi: 10.1029/2018JB017097.

Williams, S. D. P. The effect of coloured noise on the uncertainties of rates estimated from geodetic time series. *Journal of Geodesy*, 76(9-10):483–494, 2003. doi: 10.1007/s00190-002-0283-4.

Zhu, L. and Rivera, L. A. A note on the dynamic and static displacements from a point source in multilayered media. *Geophysical Journal International*, 148(3):619–627, 3 2002. doi: 10.1046/j.1365-246X.2002.01610.x.

doi:10.26443/seismica.v3i1.887

The Impact of the Three-Dimensional Structure of a Subduction Zone on Time-dependent Crustal Deformation Measured by HR-GNSS

Oluwaseun Idowu Fadugba *[1], **Valerie J. Sahakian** [1], **Diego Melgar** [1], **Arthur Rodgers** [2], **Roey Shimony** [1]

[1]Department of Earth Sciences, University of Oregon, Eugene, OR, [2]Lawrence Livermore National Laboratory, Livermore, CA

Author contributions: *Conceptualization*: O.F., V.J.S., D.M.. *Data Curation*: O.F., V.J.S., D.M., R.S.. *Formal Analysis*: O.F., V.J.S., D.M.. *Funding Acquisition*: V.J.S., D.M.. *Investigation*: All authors. *Methodology*: All authors. *Project Administration*: O.F., V.J.S., D.M.. *Resources*: O.F., V.J.S., D.M., A.R.. *Software*: D.M., A.R.. *Supervision*: V.J.S., D.M., A.R.. *Validation*: O.F., V.J.S., D.M.. *Visualization*: O.F., V.J.S., D.M.. *Writing – original draft*: O.F., V.J.S., D.M.. *Writing – review & editing*: All authors.

Abstract Accurately modeling time-dependent coseismic crustal deformation as observed on high-rate Global Navigation Satellite System (HR-GNSS) lends insight into earthquake source processes and improves local earthquake and tsunami early warning algorithms. Currently, time-dependent crustal deformation modeling relies most frequently on simplified 1D radially symmetric Earth models. However, for shallow subduction zone earthquakes, even low-frequency shaking is likely affected by the many strongly heterogeneous structures such as the subducting slab, mantle wedge, and the overlying crustal structure. We demonstrate that including 3D structure improves the estimation of key features of coseismic HR-GNSS time series, such as the peak ground displacement (PGD), the time to PGD (t_{PGD}), static displacements (SD), and waveform cross-correlation values. We computed synthetic 1D and 3D, 0.25 Hz and 0.5 Hz waveforms at HR-GNSS stations for four M7.3+ earthquakes in Japan using MudPy and SW4, respectively. From these synthetics, we computed intensity-measure residuals between the synthetic and observed GNSS waveforms. Comparing 1D and 3D residuals, we observed that the 3D simulations show better fits to the PGD and SD in the observed waveforms than the 1D simulations for both 0.25 Hz and 0.5 Hz simulations. We find that the reduction in PGD residuals in the 3D simulations is a combined effect of both shallow and deep 3D structures; hence incorporating only the upper 30 km of 3D structure will still improve the fit to the observed PGD values. Our results demonstrate that 3D simulations significantly improve models of GNSS waveform characteristics and will not only help understand the underlying processes, but also improve local tsunami warning.

Production Editor:
Gareth Funning
Handling Editor:
Mathilde Radiguet
Copy & Layout Editor:
Hannah F. Mark

Received:
April 23, 2023
Accepted:
March 10, 2024
Published:
May 3, 2024

1 Introduction

Real-time high-rate Global Navigation Satellite System (HR-GNSS) are key observational data for kinematic slip inversions (e.g., Ozawa et al., 2011; Melgar et al., 2016) that provide an important lens into large earthquake rupture physics (e.g., Melgar and Bock, 2015), as well as for real-time applications in Earthquake and Tsunami Early Warning (EEW/TEW) (e.g., Sahakian et al., 2019a). Kinematic slip inversions, traditionally based on broadband seismograms and strong motion data, are used for rapid and retrospective seismological studies to understand earthquake complexities through finite fault models, source time function and directivity, etc. (e.g., Ide, 2007; Goldberg et al., 2022). HR-GNSS waveforms are an important contribution to these models to constrain the time-dependent, low-frequency deformation of the Earth's surface. In the resulting finite fault model, this yields not only slip on the fault, but information about the rupture kinematics for each subfault, thus providing both spatial and temporal distribution of slip on a more granular level (e.g., Melgar and Bock, 2015; Melgar et al., 2020b). When performed in real-time, rapid kinematic source models are an important component of TEW approaches (hence, better rapid local tsunami modeling). HR-GNSS data also provide crucial constraints for rapid earthquake magnitude estimation, in particular for large magnitude earthquakes as the displacement metrics they provide do not saturate, unlike displacement obtained from broadband seismograms (e.g., Bock et al., 2011; Melgar et al., 2016; Sahakian et al., 2019a) and the displacement obtained from twice-integrated strong motion records do not resolve observed static offset measured by GNSS, even when high-pass filters are applied (Goldberg et al., 2021). In addition to constraining rapid finite-fault inversions for local TEW, the peak ground displacement (PGD) and time to reach peak ground displacement (t_{PGD}) as recorded by HR-GNSS can play a key role in discriminating tsunami earthquakes (TsEs) from non-TsEs (Sahakian et al., 2019a).

Static and kinematic slip inversion models using displacement time series from HR-GNSS waveforms are routinely performed using simplified 1D radially symmetric Earth models, by determining the displacement from each subfault with a 1D Green's function (Melgar and Bock, 2015), as it is oft assumed that three-

*Corresponding author: ofadugba@uoregon.edu

dimensional heterogeneities play a smaller role in the low-frequency content of waveforms than for high-frequency seismic data. However, the availability of high-rate GNSS data and the need to resolve earthquake and wave propagation details with higher frequencies and shorter wavelengths exposes the inadequacy of 1D models for analysis of large earthquake ruptures. Many studies using 1D structure observed some delays and unmodeled features in the HR-GNSS waveforms from the 2011 moment magnitude (M) 9.0 Tohoku-Oki earthquake (Yue and Lay, 2011; Melgar and Bock, 2015), as well as other earthquakes (e.g., Delouis et al., 2010).

Subduction zones present unique three-dimensional challenges that may not be well-captured by 1D models. Due to complex geometry, the resultant onshore deformation and shaking from megathrust events is likely affected by many strongly heterogeneous structures such as the slab, the wedge, the overlying crustal structure, etc. However, current models of time-dependent crustal deformation using HR-GNSS displacement waveforms, or low-frequency shaking, typically use Green's function approaches and 1D Earth structure, omitting the effects of the 3D Earth structure on the wave path, hence on the observed waveforms. In this work, we present results comparing 1D to 3D models of time-dependent crustal deformation and find that three-dimensional effects are non-negligible, and should be an important component of kinematic modeling. Although the importance of including 3D structure to model strong-motion data is well-established in the literature (i.e., Vidale and Helmberger, 1988; Olsen, 2000; Hartzell et al., 2010; Rodgers et al., 2019), this study provides quantitative estimates on the influence of neglecting 3D effects and specifically investigates the application to modeling time-dependent low frequency crustal deformation, such as that measured by HR-GNSS, still used for a variety of seismological applications.

2 Background

Previous studies have contributed to the advancement of slip models in a 3D Earth structure (e.g., Wald and Graves, 2001; Williams and Wallace, 2015; Tung and Masterlark, 2018) and show that material contrasts between continental crust and oceanic slabs have a large effect on recovering static coseismic displacements, slow slip events, slip distributions and tsunami behavior in elastic models. For example, Tung and Masterlark (2018) show that the inclusion of heterogenous crustal structure can remove nonrealistic slip artifacts in slip distributions and reduce the misfit in large seafloor displacement that contributes to prediction error of tsunami amplitudes. Williams and Wallace (2015) also show a better fit to the observed GNSS displacements by computing Green's functions using a realistically varying elastic properties with a finite element method (Aagaard et al., 2013). Hearn and Burgmann (2005) show similar effects in strike-slip settings, comparing 1D structure and homogenous half space models. They find an improvement in the estimation of the moment and centroid depth from GNSS measurements by incorporating earth's layered elastic structure in the slip inversion. This reduces the disparity between the geodetic and seismic moment estimates for large strike-slip earthquakes and suggests that time-dependent crustal deformation should be affected by depth-dependent elasticity. Langer et al. (2022) use a synthetic model of sedimentary basin to investigate the impact of 3-D elastic structure on forward models of co-seismic surface deformation and suggest the use of a layered velocity structure in static slip inversion in regions with sedimentary basins. Langer et al. (2019) show the importance of including topography in coseismic deformation modeling.

Together, these advances show that both static and dynamic (time-dependent) crustal deformation suffer from "path effects" in the same way that high-frequency ground motions as measured on strong-motion instruments do. Path effects are a common source of uncertainty in ground motion models that focus on the effects of seismic waves' path on higher frequency intensity measures (Baltay et al., 2017; Kotha et al., 2020; Landwehr et al., 2016; Zhang et al., 2022; Sahakian et al., 2019b; Kuehn and Abrahamson, 2019), and it stands to reason that they likely play a role in low to moderate frequencies (~1 Hz, that of HR-GNSS and time-dependent crustal deformation) as well.

Better modeling of time-dependent, coseismic crustal deformation can make significant contributions to improving our understanding of underlying large earthquake source processes, as well as improving warning and rapid response systems overall (Wirth et al., 2022). In this work, we show a comparison of 1D vs. 3D deterministic HR-GNSS waveforms for events in Japan to show the impact of 3D structure on accurately modeling GNSS waveforms. We choose Japan to test our hypotheses, as its seismicity, HR-GNSS recordings, and knowledge of 1D and 3D structures are ideal for our purposes. Japan has an excellent GNSS Network (~1178 stations), a good number of M7.3+ earthquakes, and both 1D and 3D velocity models (Fig. 1). We will show that the effects of including 3D structures is most important in improving the PGD at all hypocentral distances and static displacements (SD) residual values at hypocentral distances greater than 350-400 km.

3 Data and Methods

We generate 1D and 3D low-frequency synthetic GNSS waveforms of M7.3+ megathrust earthquakes in Japan and compare the 1D and 3D synthetics with the observed GNSS waveforms using several waveform intensity measures. We also test different rupture models for some of the earthquakes to investigate the effect of rupture model on the intensity measures.

3.1 Data

We focus on four M7.3+ megathrust earthquakes in Japan with good rupture models: 2011 M7.9 Ibaraki, 2011 M7.4 Iwate, 2011A M7.3 Miyagi and 2003 M8.3 Tokachi 2003 (Fig. 1; Table 1). We did not include the 2011 M9.0 Tohoku-Oki earthquake due to computational

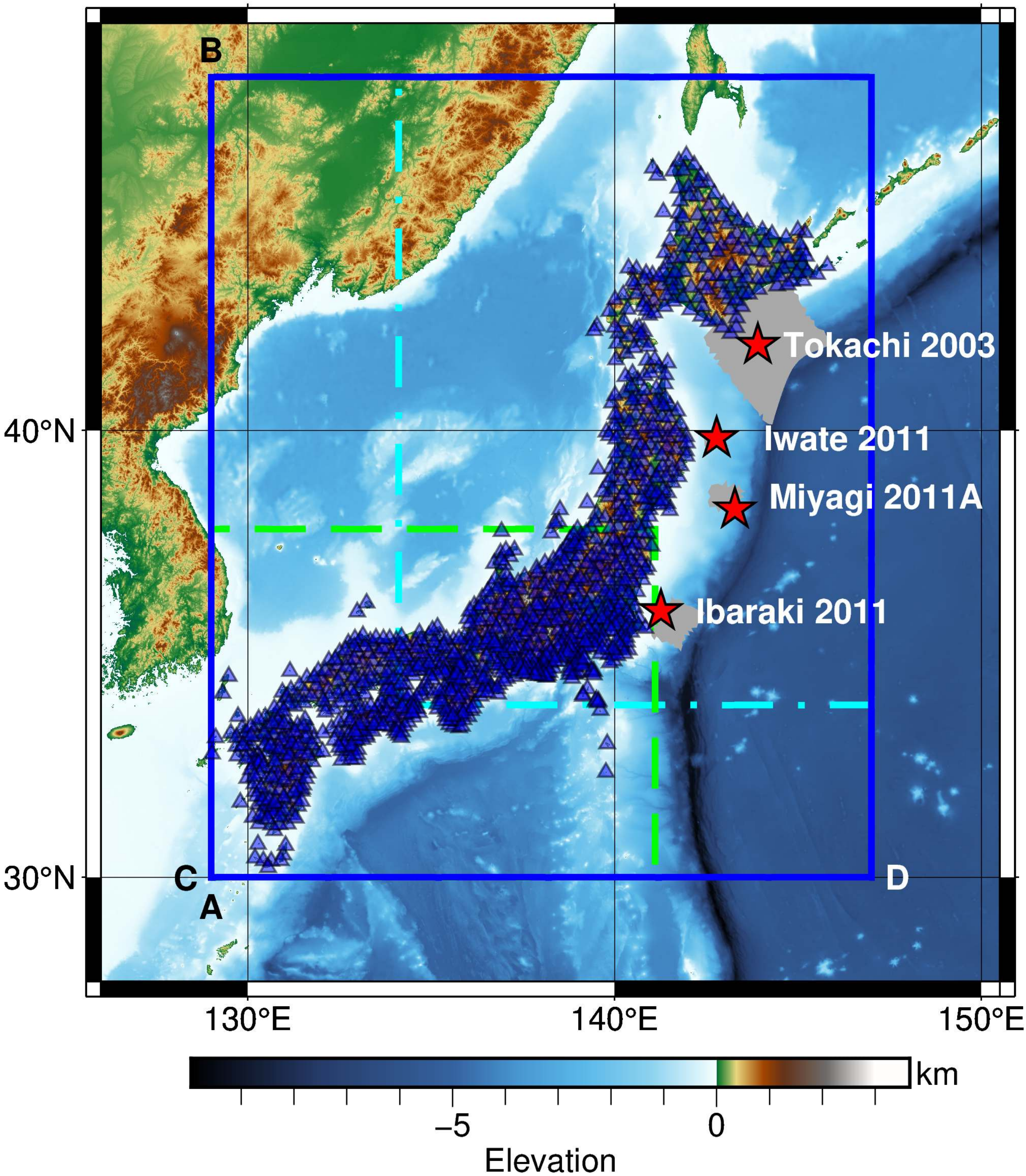


Figure 1 Study region around Japan with topography and bathymetry showing the HR-GNSS stations (blue triangles) used to observe and model at least one earthquake (SNR≥3). The figure also shows the four earthquake epicenters (red stars) used in this study and their published ruptures (the dark gray regions show the subfaults associated with the earthquakes, see Table 1). The lines show the 3D Japan Integrated Velocity Structure Model (Koketsu et al., 2008, 2009) domains: West region (green dashed line), East region (cyan dashed-dotted line) and Combined region (blue solid line). Edges AB and CD show the profile lines of the 3D Japan Integrated Velocity Structure Model presented in Fig. 4.

cost of the 3D simulations, but we expect similar conclusions with the M7.3+ earthquakes used in this study. We used the 1Hz GNSS waveforms from Ruhl et al. (2018), obtained using the Precise Point Processing approach and, from these, determine total horizontal displacement waveform $T(t)$ using Equation 1:

$$T(t) = \sqrt{N(t)^2 + E(t)^2} \tag{1}$$

We focused on the total horizontal displacement as the vertical displacement measurement in HR-GNSS are less accurate due to the distribution of GNSS satellites and generally assigned an error of about 3-5 times that of the horizontal (e.g., Geng et al., 2018; Melgar et al., 2020a). The use of $T(t)$ ensures that the more significant error in the vertical displacement compared to the horizontal displacement measurements is not influencing the 1D to 3D comparison, thus avoiding the misfit due to noise as opposed to the effect of the 3D structure.

3.2 1D Simulation Using FakeQuakes/MudPy

We used the FakeQuakes and MudPy software (Melgar et al., 2016) to generate 1D synthetic waveforms using a 1D velocity model in two steps. FakeQuakes first produces stochastic kinematic rupture models using a published rupture model as a mean slip model following the approach of Goldberg and Melgar (2020).

We give a brief description of the FakeQuakes methods, but we refer readers to Melgar et al. (2016) for more details of the method and validations. FakeQuakes generates slip distributions from the perturbations around a known slip model given a target magnitude or mean slip distribution and a prescribed fault geometry. To do so, FakeQuakes uses a von Karman correlation function to obtain the covariance matrix using correlation lengths between the subfaults associated with the rupture. The tunable parameters are the Hurst exponent and standard deviation of the slip on each subfault. It then determines the length and width of the portion of the prescribed fault geometry that will participate in the rupture using the Blaser et al. (2010) relationship which is based on the magnitude of the earthquake. It uses a lognormal probability density function approach to introduce some variability in the fault dimension. We set H to 0.4 based on Melgar and Hayes (2019) and used a uniform standard deviation of the slip (s) value of 0.9 for all subfaults. FakeQuakes then uses the Karhunen-Loéve (K-L) expansion (LeVeque et al., 2016) to determine several nonnegative slip distributions by linear combinations of the eigenmodes of a lognormal covariance matrix that are sampled from a probability density function with the desired covariance matrix. Linear combinations of more modes redistributes slip over the fault model; we set the number of modes in the K-L expansion to 72 to obtain short variability of the slip distribution necessary for kinetic rupture modeling. For example, mode 0 is roughly the alterations of the mean slip based on the lognormal covariance matrix.

To avoid an unrealistically large amount of slip, we set the limit on the peak value of slip to 40 m so that any realizations that exceed 40 m are discarded. Finally, FakeQuakes follows Graves and Pitarka (2010, 2014) to obtain the kinematic parameters of the rupture such as the rupture speed and duration of slip (rise time). The rupture speed is a factor of the shear-wave speed at the subfault depth, and rise time based on the slip at each subfault. The factor of the shear-wave speed is 0.4 in the shallow region (<10 km) and 0.8 for the deeper region (>15 km), and a linear transition in rupture speed is applied between 10 and 15 km depth. The local slip-rate function of each fault is based on the Dreger slip-rate function (Mena et al., 2010) with a fallout rate of 4.

We then use MudPy to generate displacement time series from the kinematic rupture models with an FK Green's function approach (Zhu and Rivera, 2002) using a 1D layered Earth. We used a sampling interval of 1 s and a total duration of 512 s. FakeQuakes/MudPy requires the fault and rupture models, 1D velocity model and the GNSS station locations as input parameters.

We used the Slab2.0 model (Hayes, 2018) to create a fault geometry mesh for the Japan Trench using Gmsh, a 3-D finite element mesh generator (Geuzaine and Remacle, 2009). Details of the fault files are described in the Supplementary Material (S1). We focused on the Kuril region of Japan where the M7.3+ megathrust earthquakes used in this study are located. We use the published rupture models for the four megathrust earthquakes as input mean slip distributions for FakeQuakes (Table 1). For the Ibaraki 2011 earthquake, we used the Kubo et al. (2013) rupture model (henceforth referred to as SRCMOD) and Zheng et al. (2020, henceforth referred to as Zheng) rupture model. For the Miyagi 2011A earthquake, we used the Hayes (2017, henceforth referred to as Hayes) and Zheng rupture models. For the Tokachi 2003 earthquake, we used models from Koketsu et al. (2004), Yamanaka and Kikuchi (2003), Yagi (2004) (henceforth referred to as SRCMOD, SRCMOD2 and SRCMOD3, respectively) and Hayes rupture models. We used only the Zheng rupture model for Iwate 2011 earthquake (Table 1).

The geometries of the published rupture models are planar and do not coincide with the geometry of Japan trench from Slab2.0, so, we project the slip in the rupture model for each earthquake onto the fault geometry (Fig. S2 Fadugba et al., 2023). Specifically, we project both the subfault locations of the rupture model and the centroid of the mesh of the fault geometry on a 2D plane with a strike of 210 and a dip value of 20 based on the fault geometry's general strike and dip values. We then performed linear interpolation to evaluate the strike- and dip-slip amounts from the rupture model at the mesh locations.

With FakeQuakes, we generated 100 realizations of the published rupture models using the published model as a mean model. Figure 2 shows the mean rupture model (SRCMOD) for the Ibaraki 2011 earthquake (Kubo et al., 2013) and three examples of the 100 FakeQuakes ruptures realizations from the mean rupture model. FQ Model 3 is an end-member example of a rupture model with a different slip pattern compared with the mean slip model, but with similar moment release and falling within the prescribed uncertainties of the mean slip model. The mean rupture models and examples of FakeQuakes ruptures of the other earthquakes used in this study are in the Supplementary material (Fig. S3-S4).

We adopted the 6-layer crustal velocity structure of

SN	Event Name	Origin Time (UTC)	Latitude (°)	Longitude (°)	Depth (km)	Moment magnitude (M)	Number of GNSS Stations (SNR≥3)	Rupture Models and corresponding references
1	Ibaraki 2011	2011-03-11T06:15:34	36.1083	141.2653	43.2	7.9	737	SRCMOD (Kubo et al., 2013)
2	Iwate 2011	2011-03-11T06:08:53	39.8390	142.7815	31.7	7.4	271	Zheng (Zheng et al., 2020)
3	Miyagi 2011A	2011-03-09T02:45:12	38.3285	143.2798	8.3	7.3	240	Hayes (Hayes, 2017)
4	Tokachi 2003	2003-09-25T19:50:06	41.7750	143.9040	27.0	8.3	236	Hayes (Hayes, 2017)

Table 1 Earthquakes used in this study and the corresponding rupture models. SN: Source Number.

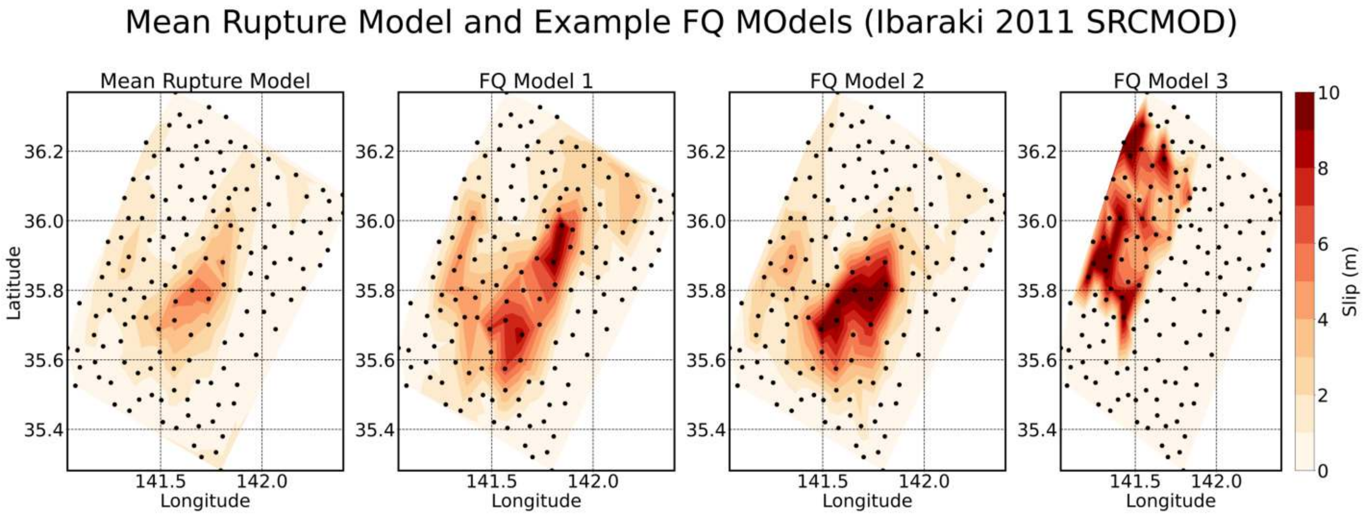


Figure 2 Mean Rupture model (SRCMOD) for Ibaraki 2011 earthquake (Kubo et al., 2013) and three examples of the 100 FakeQuakes (FQ) ruptures realizations from the mean rupture model. The color indicates the amount of slip per subfault, and the black dots signify the center of each subfault. The slip is greater overall in the FakeQuake models compared to the mean slip model in the top left to conserve the moment release in response to the change in rigidity at the subfault locations compared to the one used to generate the mean slip model. FQ Model 3 is an end-member example of a rupture model with different slip pattern compared with the mean slip model, but with similar moment release.

Hayes (2017) for all the earthquakes for its simplicity and the ease to set it up in our 1D and 3D simulations (Fig. 3). We used the isotropic Preliminary Reference Earth Model (PREM) from 40 to 200 km depth (Dziewonski and Anderson, 1981). There are other well-known 1D velocity models for Japan that could be used (e.g., Ueno et al., 2002; Hayes, 2017; Laske et al., 2013). The Japan Meteorological Agency (JMA) uses a 1D velocity model (Ueno et al., 2002) to locate earthquakes in Japan. However, the model has a series of 500 m thick layers which are less practical to setup for our 3D simulations, and we aim for consistency between the 1D and 3D simulations.

The earthquakes were recorded on a total of 1178 GNSS stations. However, to reduce computation time, we only simulate waveforms for stations with an observed total horizontal displacement signal-to-noise ratio (SNR, Equation 2) larger than or equal to 3, thus reducing the number of stations for each earthquake simulation (Table 1). SNR is defined by

$$SNR = \frac{\sigma_{signal}}{\sigma_{noise}} \tag{2}$$

where σ_{signal} is the standard deviation of 120s of recorded ground shaking after the P-wave arrival time while σ_{noise} is the standard deviation of 10s recordings before the P-wave arrival time. P-wave arrival time is defined as the origin time plus an approximate P-wave travel time (i.e., hypocentral distance between HR-GNSS station and rupture model hypocenter, divided by 6.5 km/s).

To understand the impact of the source rupture model on our synthetic waveforms, we investigated the effect of rupture models in the 1D simulations using two rupture models for the 2011 M7.9 Ibaraki and 2011A M7.3 Miyagi earthquakes and three rupture models for 2003 M8.3 Tokachi earthquake, and compare their residuals.

3.3 SW4 3D Simulations

Our 3D synthetic waveforms were computed using SW4 2.01 (Petersson and Sjögreen, 2012, 2015; Sjögreen and Petersson, 2011; Petersson and Sjögreen, 2017) published under the GPL2 license. SW4 solves the seismic wave equations in displacement formulation using a 4th order accurate summation-by-parts finite differ-

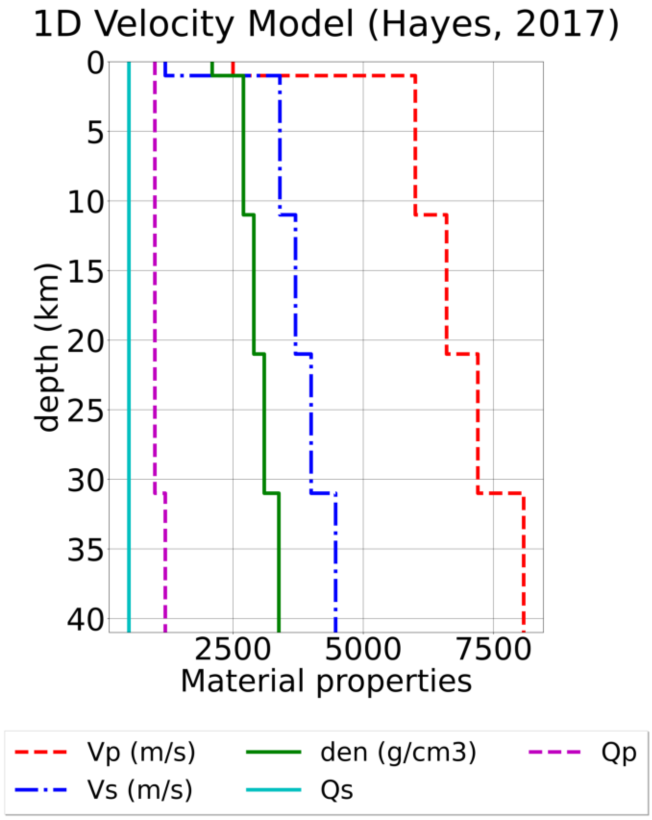


Figure 3 1D velocity model of Japan (Hayes, 2017) showing the P-wave velocity profile (red dashed line), S-wave velocity profile (blue dashed line), density profile (green solid line), and P- and S-wave quality factors (Qp and Qs) profiles (purple dashed line and cyan solid line, respectively). We used this 1D velocity model for the upper 40 km and the PREM model (Dziewonski and Anderson, 1981) from 40 km up to 200 km.

ence method, a 3D model of velocity structure, and a domain geometry that includes both topography and bathymetry. Because this process is so computationally intensive, similar to the large-scale SW4 simulations of earthquakes on the Hayward Fault (Rodgers et al., 2020), we generated simulations for all four events at both 0.25 and 0.5 Hz, to compare with observations and understand if characteristic intensity measures such as PGD require information from higher frequencies.

We used the 3D Japan Integrated Velocity Structure Model (Koketsu et al., 2008, 2009) which includes topography and bathymetry data from the ETOPO1 1 arc-minute global relief model (N.O.A.A. National Geophysical Data Center, 2009) spanning a lateral extent of latitude from 30° to 47° North (~2040 km) and longitude from 129° to 147° (~1440 km) East. We convert the 3D velocity structure from an ASCII text file to a raster file format (rfile), as it is more effective for smoothly varying 3D heterogenous structure (Fig. 4; Petersson and Sjögreen, 2017). An rfile is a binary structured grid format, and it is the most efficient and realistic method to input 3D velocity structure to SW4, hence more suitable for this study. The 3D Japan Integrated Velocity Structure Model (JIVSM) comprises 23 layers, each with constant P- and S-wave velocities (Vp and Vs), density (r) and P- and S-wave quality factors (Qp and Qs) (Table S1). The 3D structure is given in two overlapping sections (East and West Japan, Fig. 1), but were combined to create the unified 3D velocity model of Japan by extrapolating the top of each layer to regions outside the 3D structure regions following the OpenSWPC methodology (Maeda et al., 2017). The resulting rfile has 5 blocks with increasing grid spacing with depth: grid spacing of 200 m at the top to 1000 m at the bottom of the rfile. The grid sizes of rfile are independent of the grid sizes in the computational domain (Petersson and Sjögreen, 2017). The minimum grid size in the computational grid depends on the desired maximum frequency. Details of the rfile are in the Supplementary Material (S2).

Our domain depths extended from the surface (topography and bathymetry) to a maximum depth of 200 km. The maximum achievable frequency (f_{max}) is dependent on the grid size of the domain, as well as the minimum shear wave speed, as described by:

$$f_{\max} = \frac{\min V_s}{PPW \times h} \tag{3}$$

SW4 allows user to set the P- and S-wave minimum velocity values in the simulations using the global material command, thus replacing the velocity layer whose V_p and V_s are smaller than threshold values with the threshold values. We used 8 Points Per Wavelength (PPW) in the simulations and the minimum shear wave speed (minV_s) value of 1200 m/s based on the average V_S value in the upper 400 m in the 3D velocity model (Equation 3; Petersson and Sjögreen, 2017). We set the minimum P-wave velocity value in the simulations to 2500 m/s. To generate 3D synthetic waveforms with a maximum frequency of 0.25 and 0.5 Hz, we used a minimum grid spacing (h) of 600 and 300 m, respectively. We used a curvilinear mesh from the surface (topography and bathymetry) to 30-km depth with a grid spacing of 300 m and used Cartesian mesh from 30 km downwards. Within the Cartesian mesh, we applied grid refinement at 75 km depth to reduce the computational resources required for these simulations. Our grid spacing increased with depth and the associated increasing minV_s: 300 m and 600m grid spacing for the 30-75 km and 75 – 200 km depth range, respectively. For the 0.50 Hz SW4 simulations, we varied the lateral extent of the 3D domain geometry depending on location of each earthquakes (Fig. S5) to limit the maximum memory required by the simulations to ~4TB (Supplementary Material S3). We compared only the intensity measures from the common stations between the 1D and 3D simulations.

For the 0.25 Hz simulations, we selected two of the 100 rupture models from FakeQuakes for each earthquake and read them into SW4 format using the Standard Rupture Format version 2.0 source representations (Graves, 2014). We used only one rupture model for the 0.5 Hz simulations due to their high computational cost.

3.4 Comparing 1D vs. 3D Synthetic Waveforms

We compare the 1D and 3D synthetics with the observed GNSS waveforms using the total horizontal component

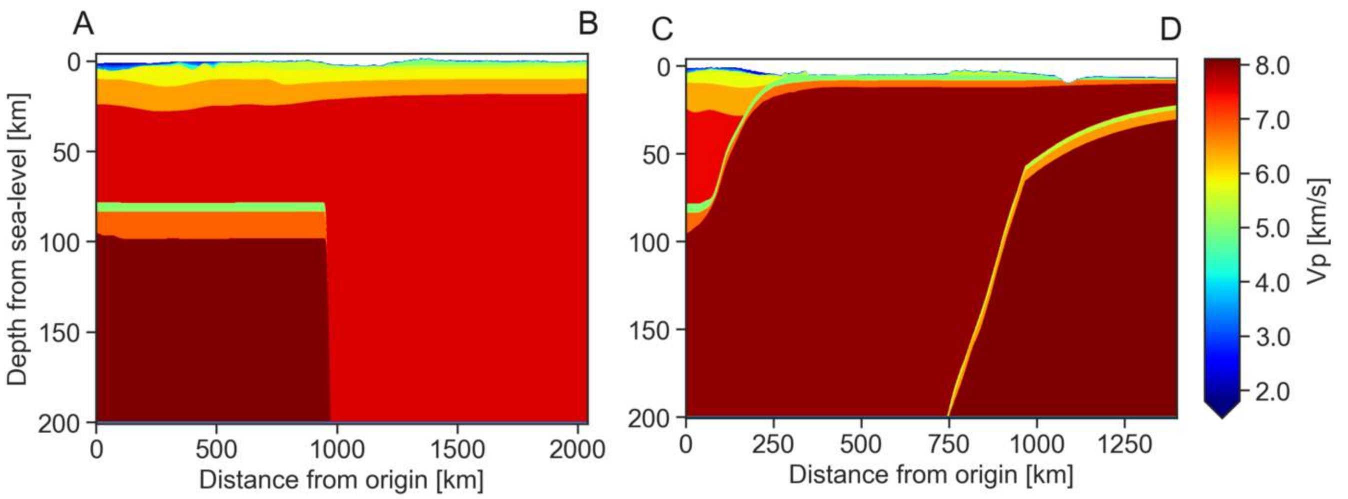


Figure 4 3D Japan Integrated Velocity Structure Model (JIVSM, Koketsu et al., 2008, 2009) shown for the AB and CD profile lines marked in Figure 1. The profile CD shows the geometries of the two subducting slabs and both profile lines best show the heterogenous velocity structure in the upper 30 km depth of the 3D velocity structure.

waveforms. In addition to wiggle-to-wiggle comparisons via waveform cross-correlation with time-shifting for both 1D and 3D synthetics, we also model the average behavior of important features of the observed waveforms over many realizations from the mean rupture models. We measure the goodness of fit by comparing the misfits of the total horizontal waveform synthetic and observed waveforms using waveform intensity measures such as the PGD as defined in Goldberg et al. (2021), t_{PGD}, and SD residuals, each described in Figure 5.

We then determine the residuals for the PGD and SD intensity measures using the equation

$$\delta_{ij,PGD} = \ln\left(\frac{PGD_{obs}}{PGD_{syn}}\right) \tag{4}$$

A residual (δ_{ij}) of 0 corresponds to perfect equivalence between observed and synthetic values, while residual values of 0.5 and 1.0 signify that the observed value is 1.6x and 2.7x the synthetic values, respectively. For the t_{PGD} residuals, we use the difference between the time it takes to reach the PGD for observed and synthetic waveforms:

$$\delta_{ij,t_{PGD}} = t_{PGD,obs} - t_{PGD,syn} \tag{5}$$

Cross-correlation values inherently compare the fit between observed and synthetic waveforms.

We investigate the variation of each intensity measure with distance by binning the intensity measures with respect to the hypocentral distance, the distance between rupture model hypocenter and the HR-GNSS station. Each intensity measure is defined for a paired station and rupture model (Equations 4, 5), and we combined the residuals from all stations and rupture models into a single dataset and binned with respect to the hypocentral distance. The residuals in each bin are plotted using box and whisker plots. These combine the minimum and maximum values with the quartiles into one useful graph. It consists of a horizontal line, drawn according to scale, and a box drawn from the lower (Q_1) to upper (Q_3) quartile with a vertical line marking the median. The minimum and maximum values of the whisker correspond to the smallest and largest data points from the dataset that fall within 1.5 times the inter-quartile range (IQR = Q_3-Q_1). Outliers are observed data points that are more than 1.5 times the IQR below Q_1 or more than 1.5 times the IQR above Q_3. For a normal distribution, the IQR contains 50% of the population and 1.5 of the IQR contains about 99%. We removed the outliers outside the whiskers to improve readability.

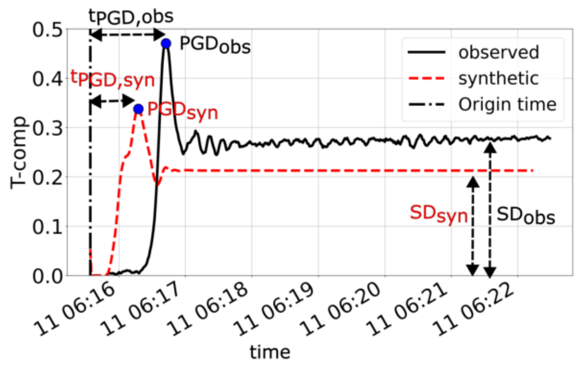


Figure 5 Schematic representation of the definitions of the intensity measure used in comparing the total horizontal waveform synthetic ("syn") and observed ("obs") waveforms. Black solid and red dashed lines are the observed and synthetic waveforms, respectively. The blue dots show the Peak Ground Displacement (PGD) for the observed and synthetic waveforms and their corresponding time to reach the PGD (t_{PGD}). The figure also shows the definition of static displacement (SD). The amplitude and time axes values in this figure are arbitrary.

4 Results and Discussions

4.1 Comparing 1D and 3D Residuals

First, as a control, we study the impacts of varying only the source model by investigating the residuals using only 1D velocity structure, with different published rupture models. The PGD residuals ($\delta_{ij,PGD}$, Equation 4) for all the earthquakes in the 1D simulations increase with distance but are generally below 2 (Fig. 6). On each boxplot, residuals for each model are shown as patterned box and whisker plots, including blue circle patterns (SRCMOD), orange grid (SRMOD 2), light blue circled (SRCMOD 3), gray slanted (Hayes), and orange diamonds (Zheng). The red horizontal line represents the zero residual line. In the 1D simulations, PGD residuals do not change significantly with distance when we used different rupture models for the same earthquake (e.g., Hayes, SRCMOD, SRCMOD2, and SRCMOD3 for the 2003 Tokachi earthquake). Therefore, any deviations in the PGD residual for the same rupture model in 3D simulations are most likely due to the 3D Earth structure. We observed that the PGD residuals for MudPy 1D Zheng model is lower than that of the MudPy 1D SRCMOD model, but we will show later that the residuals for 3D velocity models are still lower than the corresponding 1D models.

We evaluated a possible bias in the choice of the 1D velocity model since the mean rupture models (Table 1) are derived from 1D crustal models by other researchers (e.g., Zheng et al., 2020). We have shown that the choice of 1D velocity model, even though different from the source models' 1D model, do not affect the conclusions and the PGD residuals in the 1D simulations will still very different from the 3D simulations (Fig. S6). We performed 1D simulations for the Ibaraki 2011 earthquake using the SRCMOD mean rupture model but using 1D velocity models used by Koketsu et al. (2004) and Zheng et al. (2020). We compared the PGD residuals of the resulting waveforms using these 1D velocity models with respect to the observed waveforms. The comparison plot shows that the PGD residuals using these additional 1D velocity models are different in some sense but are not significantly different compared to the trend of residuals observed for the 3D simulations shown later. Therefore, any deviation from the PGD residual in the 1D simulations is due to the path rather than the choice of the 1D velocity model.

The effect of this is exemplified in the comparison of the observed and the MudPy 1D and SW4 0.25 Hz and 0.5 Hz synthetic waveforms at stations 0041 and 0043 for the Ibaraki 2011 earthquake for one of the 100 FakeQuake ruptures using the SRCMOD mean rupture model (Kubo et al., 2013) (Fig. 7). The MudPy 1D waveforms are very simple, but the SW4 waveforms better capture the variability in the observed waveforms. Specifically, the 3D waveforms and in particular the higher frequency 3D waveforms better capture the dynamic shaking in addition to the static offset observed at each station. This includes capturing a commonly observed dynamic overshoot, such as that observed in the North component of stations 0041 and 0043 for the Ibaraki earthquake, at ~60s (Fig. 7).

In a map view, we further show the effect of the 3D structure by overlaying the magnitude of velocity waveforms at the surface as a function of time on a topography/bathymetry map to highlight the spatial and temporal variation of the wavefront as it propagates (Fig. 8). The wavefronts appear spherical up to about 120s (Fig. 8E) and reveal a strong energy propagating SE away from the land. At 140s (Fig. 8F), the wavefronts show evidence of a waveguide on the low-velocity wedge as the energy propagates at a lower velocity within the wedge area exemplifying the effect of the 3D structure. The extent of the packet of energy coincides with the geometry of the Japan trench. The packet of energy within the wedge continues to propagate northward as the wavefront propagates through Japan. At 200s (Fig. 8H), the wavefronts reveal a basin effect in the Nankai and Sagami Troughs located SW of the Japan Trench and in the Sea of Japan. The wavefront also shows a waveguide phenomenon in the low-velocity wedge of the Nankai Trough and the packet of energy propagates westward at a slower velocity in the wedge even though the earthquake is located on the Japan Trench. From 270s onwards (Fig. 8J), the wavefront traveling northward through the wedge appears to bifurcate into the bay region towards Tomakomai and the other energy continues northward within the wedge. The observed waveguiding in the shallow slabs and the wave amplification in the Nankai and Sagami Troughs area show that lower frequencies still demonstrate non-negligible path effects, which may be important to the seismic hazard of Japan. Furthermore, this demonstrates that three-dimensional effects are important to include in kinematic slip models, as they may currently be wrapped into the source model.

This observation is distinctly different from the subduction guided waves observed from deep earthquakes on the subducting Pacific plate in Japan (Furumura and Kennett, 2005), as well as other regions globally (Furumura and Kennett, 1998; Furumura and Singh, 2002; Sahakian et al., 2018; Mann and Abers, 2019). In Japan, Furumura and Kennett (2005) observed an anomalously large intensity on the eastern seaboard of northern Japan from deep-seated earthquakes and the waveforms show a low-frequency (f<0.25 Hz) onset for both P and S waves, followed by large, high frequency (f>2Hz) later arrivals with a long coda. They did not observe the characteristics of frequency-selective wave propagation for subduction zone earthquakes with hypocenter depth less than 185 km. They explained this observation as arising from scattering of seismic waves by an elongated scatterer parallel to the plate margin. Despite the similarity in the phenomenon, the Ibaraki 2011 earthquake shown in Figure 8 has a hypocenter depth of 43.2 km and the maximum frequency in the waveforms is 0.25 Hz. We observed that the intense shaking is concentrated within the shelf regions and is bounded by the trench geometry. This shows that the shaking may be due to waveguide phenomena within the low-velocity wedge. Indeed, in other subduction zones such as the Hikurangi, the sedimentary wedge is demonstrated to act as a waveguide, increasing shaking

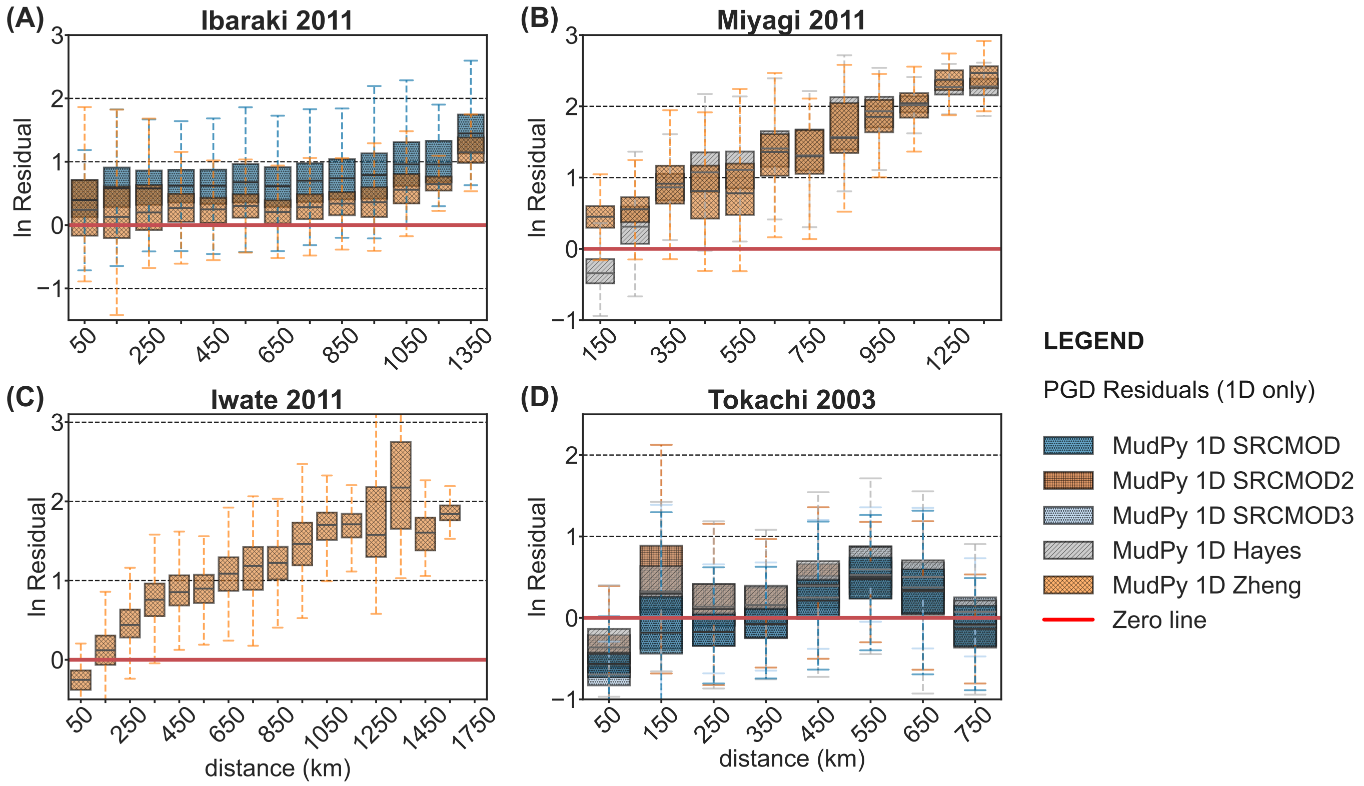


Figure 6 PGD Residuals using only the 1D velocity model, demonstrating the effects of varying solely the source model for all the four earthquakes using all the 100 random realizations of the mean rupture model. A, B and D show the effect of rupture models on the PGD residuals from 1D simulations using Ibaraki 2011, Miyagi 2011A and Tokachi 2003 earthquakes. On each boxplot, residuals for each model are shown as patterned box and whisker plots, including blue boxplot with circle patterns (SRCMOD), orange boxplot with grid patterns (SRMOD 2), light blue boxplot with circled patterns (SRCMOD 3), gray boxplot slanted patterns (Hayes), and orange boxplot with diamonds patterns (Zheng). The red horizontal line represents the zero residual line.

and dynamic stresses for longer period ground motions (Wallace et al., 2017; Kaneko et al., 2019). A more detailed examination of the waveguide is beyond the scope of this paper.

The PGD, t_{PGD}, SD residuals, and cross correlation residual maps for the Ibaraki 2011 earthquake showing the spatial variation of the residuals are in the supplementary materials (Fig. S8). The PGD residual is generally near zero and positive, but $<ln(1)$. We observed the residual is more positive near the coastal region of the Nankai Trough. However, an isolated zone with negative PGD residuals is observed near Kanazawa at the Japan Sea Margin. The t_{PGD} residual is generally positive and below 50 s, but slightly negative on the Japan Sea Margin. It is also noteworthy that the static displacement residuals are generally near zero but become more variable farther away from the hypocenter, especially toward SW Japan. The cross-correlation values between the SW4 3D waveforms and the observed waveforms show a decay in values with distance but are generally above 0.7.

4.2 Residual Analyses

Comparing 1D and 3D residuals, we observed that the 3D simulations residuals are clearly near zero (closer to the observed intensity measures) than the 1D residuals at all distances, except for the Tokachi earthquake (Fig. 9 and S9). The distributions of the intensity measures show improved fitting to the observed waveforms in the 3D simulations. These results suggest that accounting for path-specific 3D structure improves the fit to the observed waveforms compared to the 1D simulations. The width of the residual distributions is to some degree controlled by the parameters used to vary the random slip model realizations upon a mean model as described in the methods; however, there is not necessarily a one-to-one relationship between these parameters (such as h) and the width of the residual distributions here. Furthermore, the difference between 1D and 3D residuals is significantly greater than the difference between residuals for any given models of an event (Figure 6), demonstrating that the structure has a greater effect than any potential bias due to the source model selection.

For a more quantitative aggregate comparison, we determine the difference between the magnitude of the 3D median residuals compared to 1D median residuals for each residual boxplot (Equation 6, Fig. 10). We compute the difference as:

$$\delta_{|3D|-|1D|} = |\delta_{3D}| - |\delta_{1D}| \quad (6)$$

where $|\delta_{1D}|$ and $|\delta_{3D}|$ are absolute values of the 1D and 3D median residuals, respectively. The median residual difference measures how much the 3D median residual is closer to the zero value (i.e., fits the observed

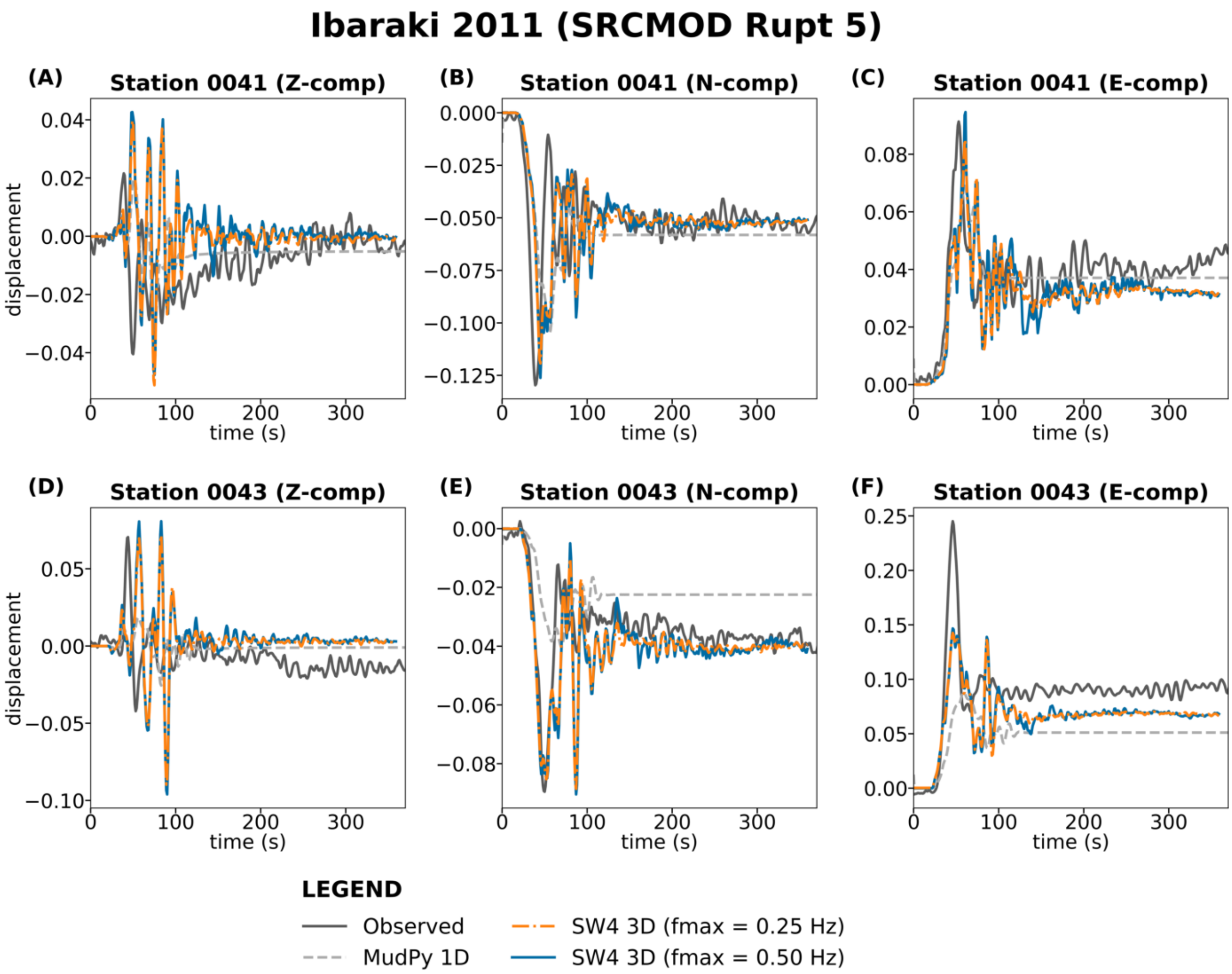


Figure 7 Comparing the observed (dark gray solid line) and three synthetic waveforms: MudPy 1D (dashed gray line), SW4 0.25 Hz (dashed-dotted orange line) and 0.50 HZ (blue solid line) waveforms at stations 0041 and 0043, respectively. The MudPy 1D waveforms are very simple, but the SW4 waveforms better capture the variability in the observed waveforms. The observed waveforms were shifted back by 20 s to fit the synthetic waveforms. The figure shows the vertical (Z-comp) and the horizontal components (N-comp and E-comp) of the waveforms.

waveform) than the 1D simulations. A negative value of median residual difference shows that 3D simulations fit better to the observed intensity measure than the 1D simulations and vice versa. Note that this convention is different for the median residual difference for the cross-correlation values because a positive median residual difference for the cross-correlation shows that the 3D simulations fit the observed waveforms better.

To determine if the 1D and 3D residuals are statistically different from each other (i.e., come from different distributions), we perform Kolmogorov-Smirnov (K-S) tests (Kolmogorov, 1933; Smirnov, 1948) on the 1D and 3D residuals for each earthquake. Two distributions are significantly different when the statistical value (KS-stat) is above a critical value which is a function of the number of samples of each distribution, and when the p-value is below the significance level of 0.05.

Considering the variation of the median residual of the intensity measures with distance, 3D simulations consistently have lower PGD median residuals (Fig. 10A) for all simulations with statistical significance (Fig. 10B), except for Tokachi 2003 (Fig. 10A). The t_{PGD} median residuals are consistently lower in the 3D simulations, generally between 250 and 700 km hypocentral distance except for the Ibaraki 2011 earthquake (Fig. 10C). The static-displacement median residuals are similar up to about 400 or 500 km (i.e., near zero), but the 3D simulations fit the observed static displacement better at longer distances (i.e., negative) (Fig. 10E). The cross-correlation median values are slightly higher in the 3D simulations, especially above distances of about 300 km, excepting the Ibaraki 2011 SRCMOD earthquake (Fig. 10G); however, we have no explanation for why this particular model shows lower cross-correlations other than that it may be related to the source inversion parameters.

The plots of the p-value with distance for all the earthquakes show that the 1D vs 3D intensity measure residuals vary with distance for all simulations (Fig. 10). Specifically, the K-S tests show that the distributions of

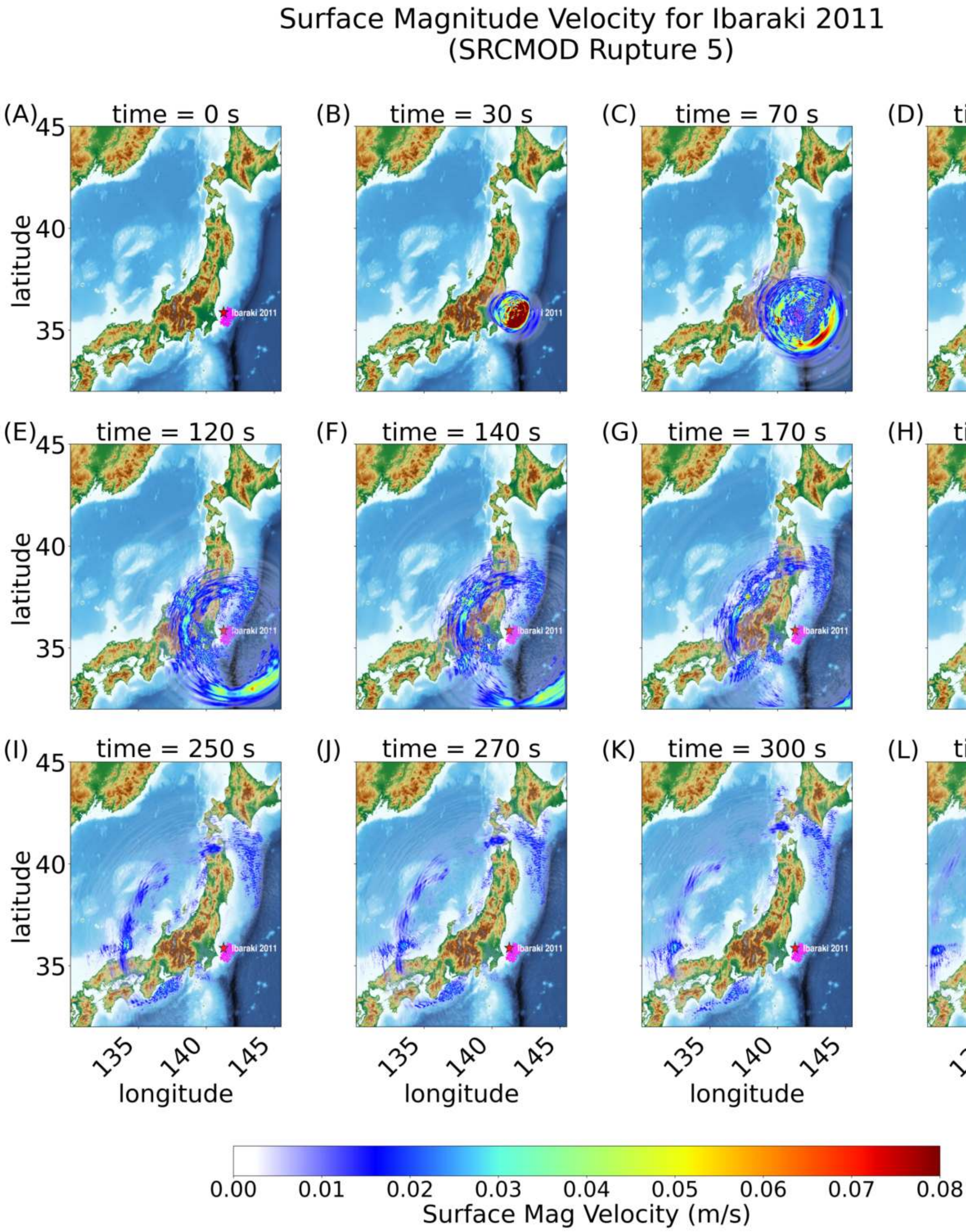


Figure 8 Waveform propagation of Ibaraki 2011 earthquake using rupture 5 of the 100 FakeQuakes random realizations of the SRCMOD mean rupture model (Kubo et al., 2013), showing the effect of 3D velocity structure. The maximum frequency of the simulation is 0.25 Hz. The rupturing subfaults are shown as pink grid cells, and hypocenter as a star. Color bar shows the surface magnitude velocity in m/s.

the PGD residuals in the 1D and 3D simulations are significantly different for all simulations up to hypocentral distance of 1000 km, and below 700 km for the Miyagi 2011 simulations (Fig. 10B). The t_{PGD} residual distributions are significantly different below 600 km distance, except for Ibaraki 2011 and Tokachi 2003 earthquakes below 400 km, which corresponds to the distance range where there is a better fit in PGD residuals for the 3D simulation (Fig. 10D). For both PGD and t_{PGD} residuals, the numbers of samples are generally smaller where the distributions are not significantly different. Conversely, the p-value plots for the static displacement and cross correlation show similar distributions between the 1D and 3D residuals (Fig. 10F and H).

The observed consistent overall increase in 1D and 3D residuals with distance may be because the source rupture model was derived with a 1D Green's function. The general trends in the PGD residuals show the 1D and 3D synthetic amplitudes generally decay faster than the observed amplitudes with distance, suggesting the

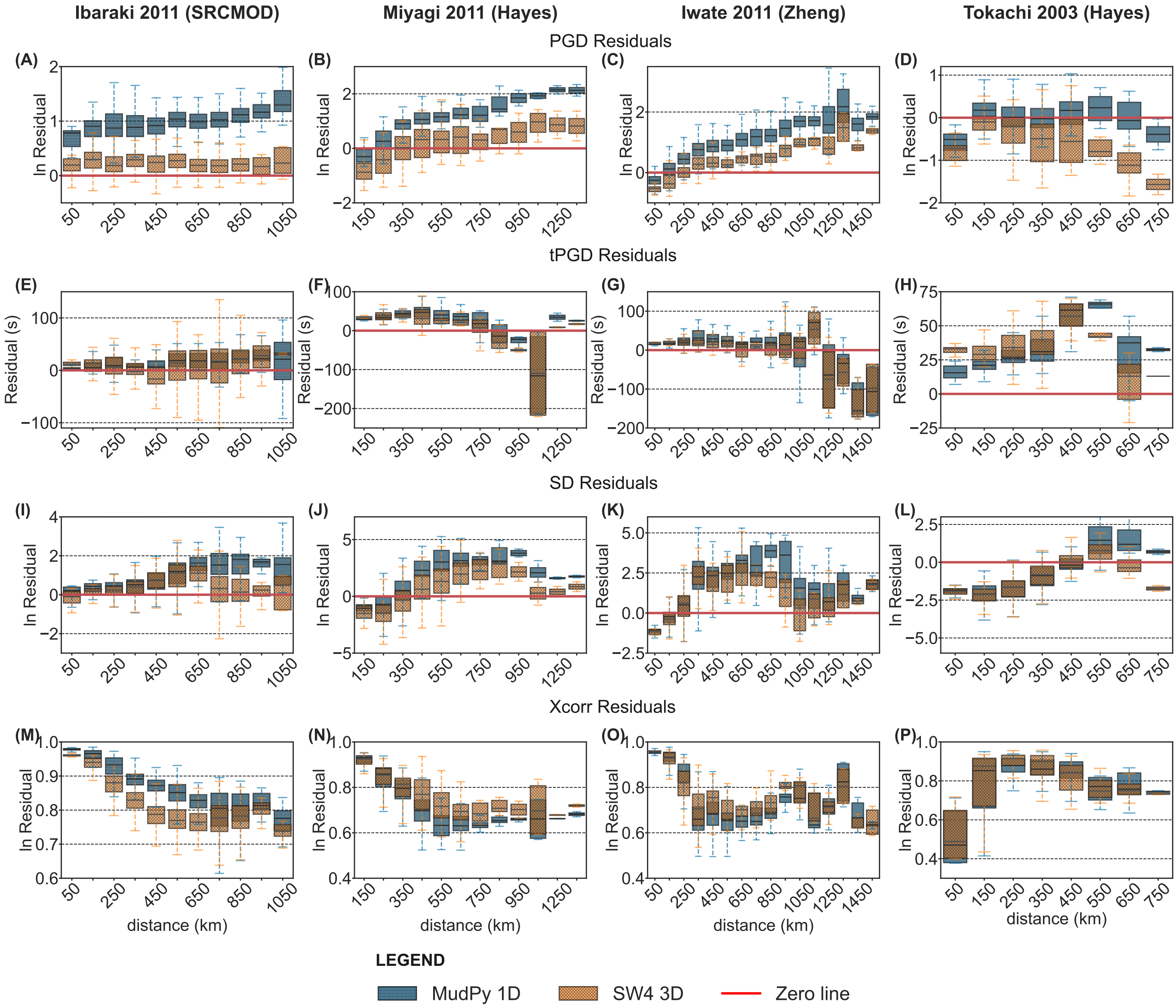


Figure 9 Comparing MudPy 1D vs SW4 3D residuals between the synthetic to observed GNSS waveforms for Ibaraki 2011 (SRCMOD), Miyagi 2011 (Hayes), Iwate 2011 (Zheng) and Tokachi 2003 (Hayes rupture model) with fmax = 0.25 Hz. (A-D) PGD residuals, (E-H) t_{PGD} (s) residuals, (I-L) static displacement residuals and (M-P) cross correlation values. We compare only the residuals of two corresponding rupture models in the MudPy and SW4 synthetic simulations. The blue boxplots with circle hatched filling represents the MudPy 1D residuals while the orange boxplot (diamond hatch style) represents the SW4 3D simulation. The red horizontal line represents the zero residual line.

variation in attenuation values within a layer unit in the 3D earth structure. Evaluation of the effect of the 1D velocity-derived rupture model and possible variation of attenuation within a layer on the general trend is beyond the scope of this study.

4.3 General Intensity Measure Residuals for Each Earthquake

The intensity measures for each simulation without considering the variation with distance show that the 1D vs. 3D residual distributions are significantly different for all simulations and there is a general reduction in the median residual values (hence, a better fit) in the 3D simulations compared to 1D simulations (Fig. 11).

Of greatest significance, we observed that the PGD residuals in the 3D simulations are smaller by about 0.4 -0.6 units compared to 1D simulations for all simulations, except for Tokachi 2003 (Hayes and SRCMOD3) models. Also, the t_{PGD} in the 3D simulations, in general, better fit the observed than 1D simulations by about 4 seconds, except for Ibaraki 2011 (Zheng) and Tokachi 2003 earthquake simulations. There is a slight reduction in the median static displacement residuals in the 3D simulations except for Ibaraki 2011 (Zheng) and Miyagi 2011 (Hayes) simulations. The 3D simulations generally have higher median cross-correlation values than 1D simulations, up to about 0.03. These results demonstrate that 3D structure plays a large, and statistically significant, role in accurately modeling the PGD and SD, as well as time-dependent characteristics of displacement time series (Fig. 11).

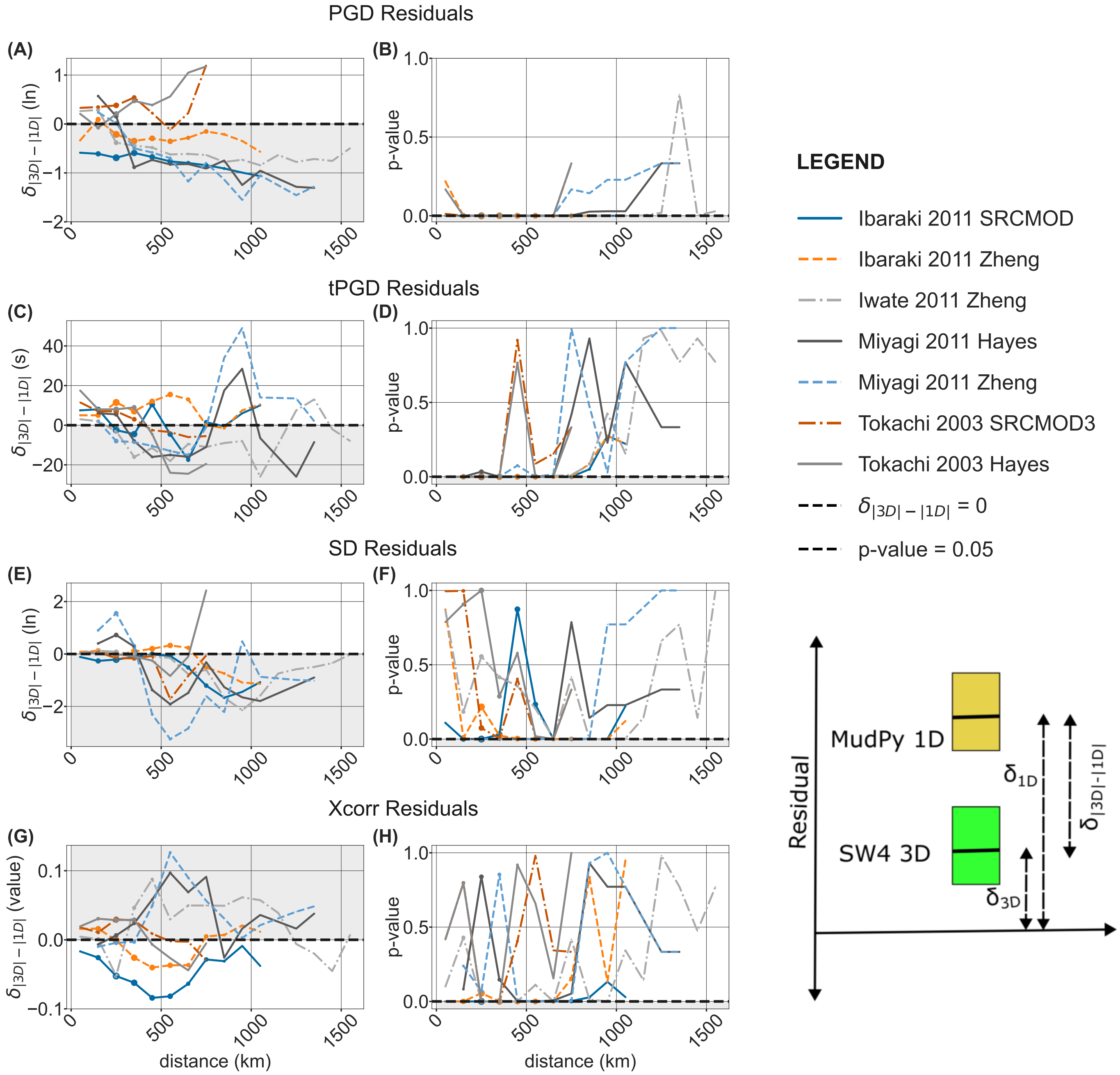


Figure 10 Median residual difference and the P-value for the PGD, t_{PGD}, SD residuals and cross correlation values for all the simulations. Blue solid lines: Ibaraki 2011 SRCMOD, orange dashed lines: Ibaraki 2011 Zheng, gray dashed-dotted line: Iwate 2011 Zheng, Black solid lines: Miyagi 2011 Hayes, blue dashed lines: Miyagi 2011 Zheng, orange dashed-dotted lines: Tokachi 2003 SRCMO3, gray solid lines: Tokachi 2003 Hayes simulations. The gray shaded regions in (A), (C), (E) and (G) represent regions where "3D fits better than 1D" while the white regions represent "1D fits better than 3D". The gray shaded regions in (B), (D), (F) and (H) represent regions where 1D and 3D residuals are statistically different from each other (i.e., come from different distributions) while the white shaded regions represent regions where 1D and 3D are from the same distribution. The bottom right schematic is a visual representation of the meaning of the mean residual difference.

4.4 Effect of 3D Structure in the Upper 0-30 km

To understand if a well-constrained shallow structure plays a larger role than deeper structure in accurately modeling time-dependent crustal deformation from mid-crustal earthquakes, we tested the effect of 3D structure in the upper 0-30 km on our simulations using the Ibaraki 2011 and Miyagi 2011 earthquakes as case studies. The Ibaraki 2011 earthquake is located at 43.2 km which is below the upper 0-30 km, while the Miyagi earthquake has a focal depth of 8.3 km, so the earthquake is within the 0-30 km structure (Fig. 1 and 12). We used the upper 0-30 km of the 3D structure because it is the depth region where we observed the most lateral structural heterogeneity.

To do this, we created another rfile for a 3D velocity model involving only the upper 0-30 km depth of the unified 3D velocity model of Japan, which is an extrapolated version of the 3D Japan Integrated Velocity Structure Model (Koketsu et al., 2008, 2009). The SW4 simu-

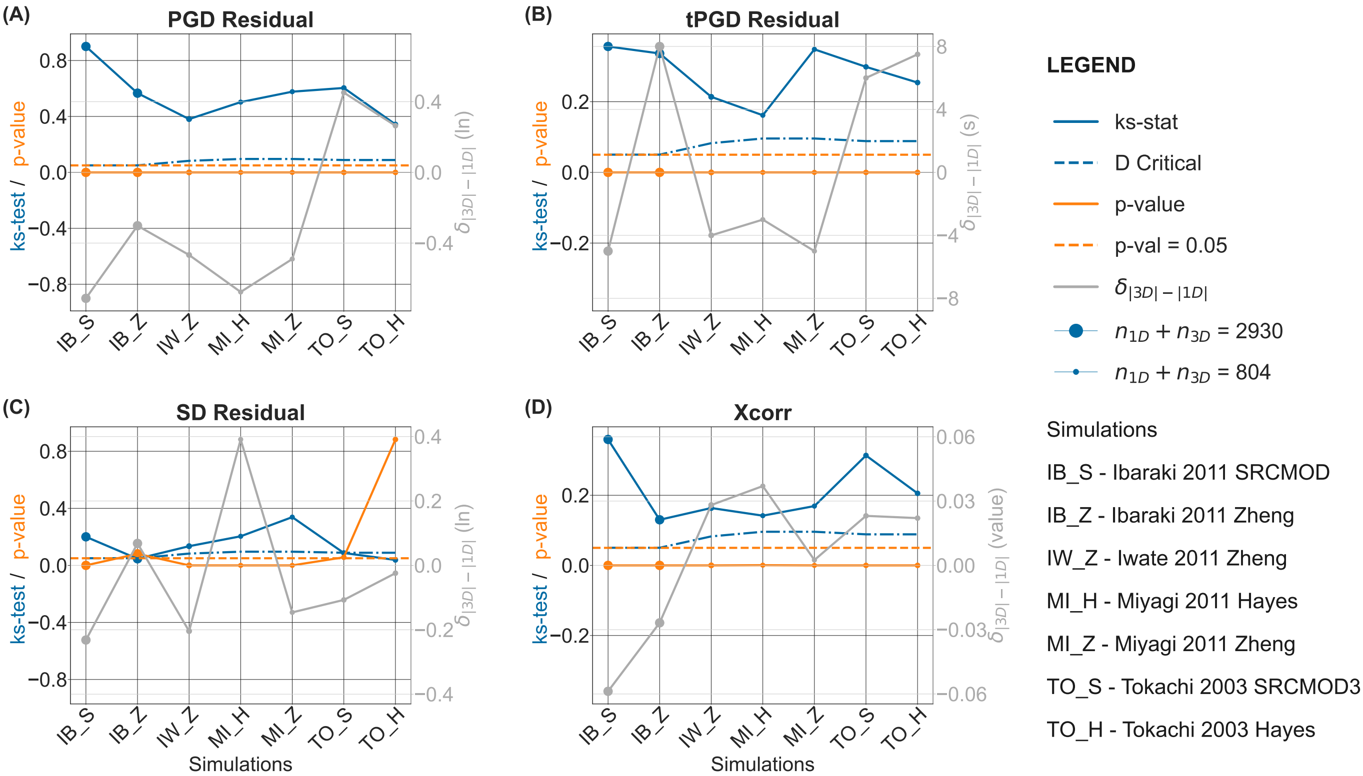


Figure 11 K-S test results and median residual difference between entire MudPy 1D and SW4 3D residual distributions for all simulations. The points are sized by the population size in that bin. Blue solid lines represent the K-S statistic value, blue dash-dotted lines represent the critical value, orange solid lines represent the p value, orange dashed lines represent the p value line of 0.05, and the gray solid line represents the median residual difference of zero.

lation is setup to use the 3D structure up to 30 km depth and a 1D velocity model, similar to the MudPy 1D simulations from 30 km to 200 km depth.

The residuals for the Ibaraki 2011 and Miyagi 2011 earthquake simulations involving 30km-depth 3D structure (3D_30km) and the 200km-depth 3D structures (3D_200km) are consistently lower than residuals from a purely 1D simulations without any 3D structure (Fig. 12). Comparing the two SW4 simulations to the MudPy 1D simulation reveals that the residual values from the 3D_30km simulation are similar to the residuals from the 3D_200km simulation up to a hypocentral distance of about 600 km. However, the residual using the 3D_200km simulation is smaller (i.e., better fit) than the 3D_30km simulation above the 600 km distance. The static displacement residuals are similar at all distances.

This result shows that the reduction of the PGD residuals in the 3D simulations is a combined effect of both shallow and deep 3D structures at hypocentral distances >~600 km. Hence, incorporating only the upper 30 km of a 3D structure will still improve the fit to the observed PGD values compared to purely 1D simulations, especially in regions where a deep 3D structure is not available. In other words, the 30km-depth structure plays a role in reducing the PGD residuals, but since the PGD residual compared to the observed waveforms is further reduced in the 3D_200km simulation for larger hypocentral distances, the deeper structure still contributes to the lower residuals. This result is important both in understanding what scale of structure should be included in 3D models, but also in estimating the computational demand in accurately modeling these time series.

4.5 Effect of Maximum Frequency on the Waveform Intensity Measures

Another important question is whether the reduction in the residuals between observed and SW4 3D simulations will persists at higher maximum frequency. To answer the question, we generated 0.50 Hz synthetic waveforms for all the four earthquakes using SW4. We varied the lateral extent of the 3D domain geometry depending on location of each earthquake, thus including fewer stations (Fig. S5), and used one of the 100 ruptures from the FakeQuakes realizations of the mean rupture models to reduce computational cost (Supplementary Material S3). We compared only the intensity measures from the common stations between the 1D and 3D simulations.

For the Ibaraki 2011 earthquakes, we observed similar trends in the PGD, t_{PGD}, SD residuals, and cross correlation values compared to the 0.25 Hz SW4 3D simulation (Fig. 13). However, the median residual difference in the PGD residual compared to the MudPy 1D simulation shows a consistent further reduction in the 0.5 Hz simulation. Hence, even though the overall trend in the residuals persists between the 0.25 Hz and 0.50 Hz simulations, the 0.50 Hz better fits the observed waveforms. Figure 14 shows the PGD residuals for the other earthquakes and rupture models. The figure further validates the reduction in the residual in the 0.50 Hz simulations, except for the Tokachi earthquake (Fig. 14).

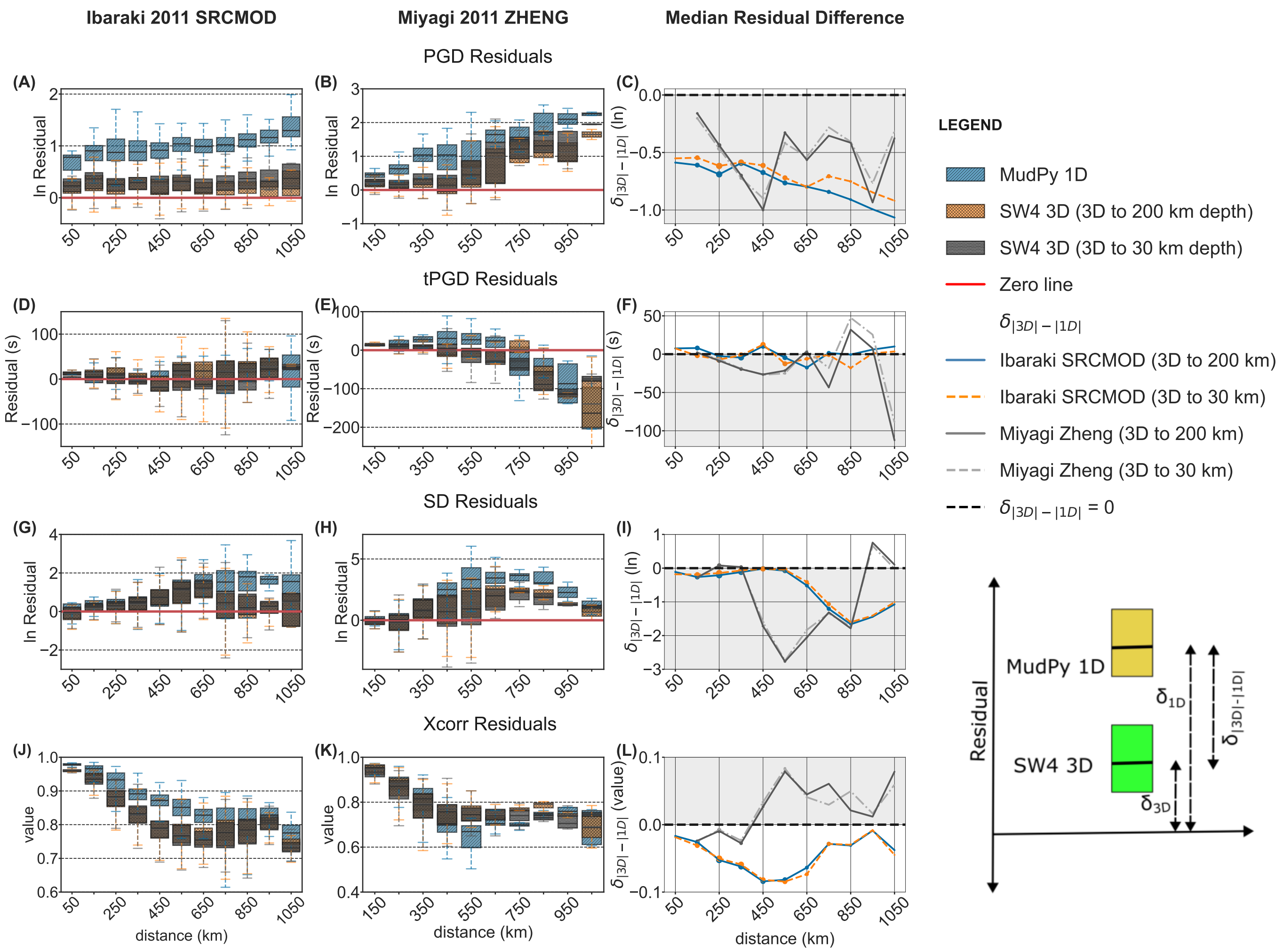


Figure 12 Effect of 3D structure in the upper 0-30 km depth on PGD, t_{PGD}, SD residuals and cross correlation values for Ibaraki 2011 and Miyagi 2011A earthquakes. We compare the MudPy 1D and SW4 3D residuals using 200 km- and 30 km-3D structure at different hypocentral distances. Ibaraki 2011 earthquake is located at 43.2 km depth while Miyagi 2011A earthquake is located at 8.3 km depth, so it is located within the upper 0-30 km depth. The figures on the right column show the median residual difference for the PGD, t_{PGD}, SD residuals and cross correlation values for two simulations compared to the MudPy residuals. Blue boxplots (slant lines hatched style): MudPy 1D residuals; orange boxplots (diamond hatch style): SW4 3D simulation using 200 km-3D structure; gray boxplots (circle hatched style); SW4 3D simulation with 3D structure up to 30 km depth; red horizontal line: zero residual line; blue solid lines: median residual difference for the simulation using the 200 km-3D structure; orange solid line: median residual difference for the 30-km-3D structure; gray shaded regions in left column: regions where "3D fits better than 1D"; white regions represent "1D fits better than 3D". Bottom right schematic is a visual representation of the meaning of the residuals presented here.

5 Conclusions and Future Work

We present 1D and 3D simulations of four M7.3+ earthquakes in Japan and showed the need to include realistic 3D structure with modern computational approaches and avoid the oversimplification of 1D GNSS models. In the 1D simulations, using different rupture models, PGD residuals do not change significantly with distance for the same earthquake. Therefore, any deviations in the PGD residual for the same rupture model in 3D simulations reveal the effect of the 3D structure. Comparing 1D and 3D residuals, we observed that 3D simulations show improved fits to the observed waveforms, demonstrating that the unmodeled waveform in the 1D simulation is due to the structure (path). However, the observed overall trends in 1D and 3D residuals with distance are likely related to a source model derived with the assumption of 1D structure or the variation of attenuation parameters within each layer in the 3D structure.

PGD median residuals with distance show that 3D simulations consistently have lower residuals for all simulations, except for Tokachi 2003. The t_{PGD} median residuals are consistently closer to zero for the 3D simulations, generally between 250 km and 700 km distance and up to 1150 km for Iwate 2011 earthquake. The SD median residuals are similar in both 1D and 3D simulations up to about 400 or 500 km, but the 3D simulations fit better at greater distances. The cross-correlation median values are slightly higher in the 3D simulation above hypocentral distance of about 300 km, except for the Ibaraki 2011 earthquake. The K-S tests show that the distributions of the PGD residuals in the 1D and 3D simulations are significantly different for all simulations up to 1000 km distance and 800 km for the Iwate

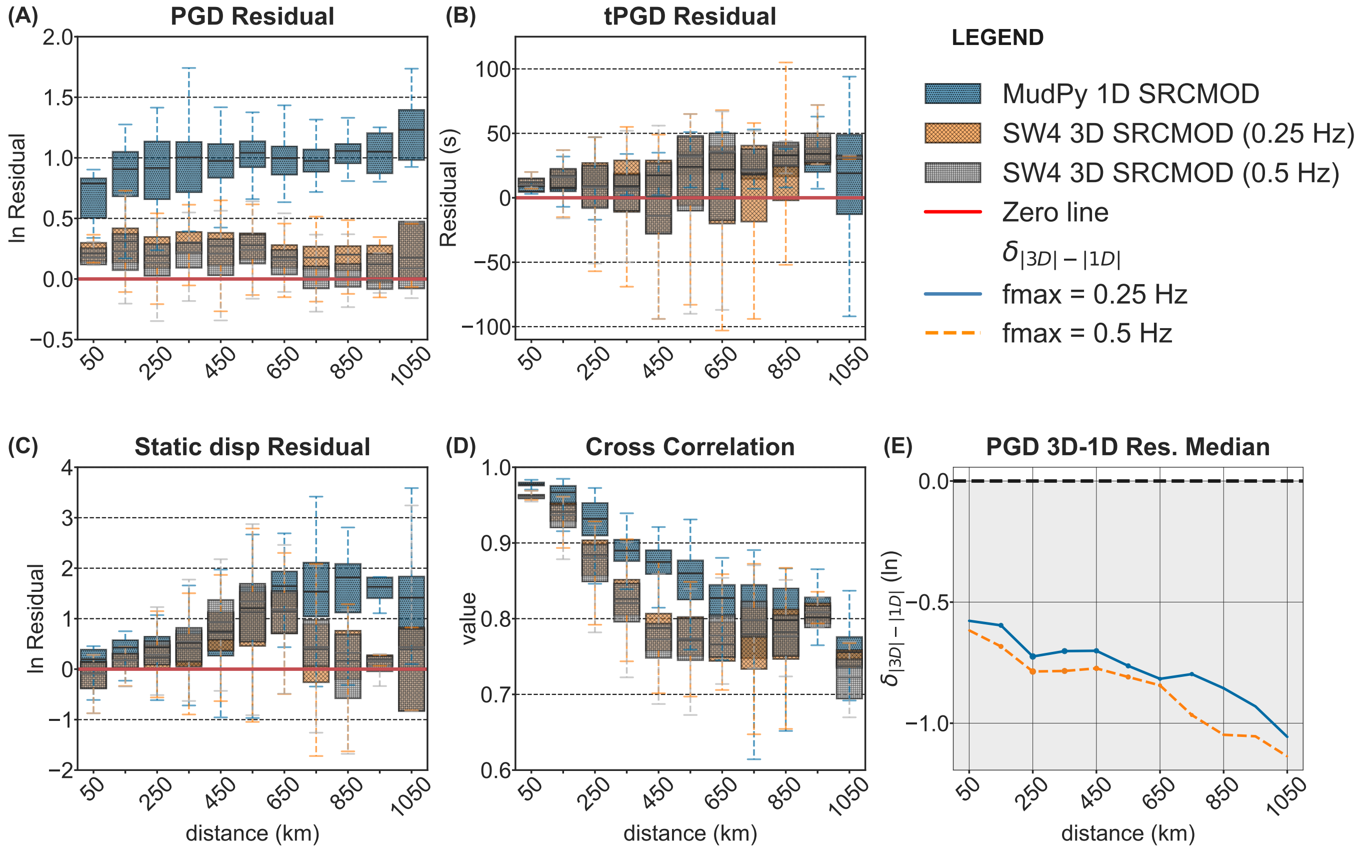


Figure 13 Effect of the maximum frequency (fmax) on PGD, t_{PGD}, SD residuals and cross correlation values for Ibaraki 2011 earthquake (using rupture 5 with SRCMOD mean rupture model). Blue boxplots (circle hatched style): MudPy 1D residuals; orange boxplots (diamond hatch style): SW4 3D simulation with fmax of 0.25 Hz; gray boxplots (crossed hatched style): SW4 3D simulation with fmax of 0.50 Hz; red horizontal line: zero residual line. Figure (E) shows the median residual difference for the PGD residuals compared to the MudPy residuals. Blue solid and orange dash lines represent the median residual difference for the simulation with fmax of 0.25 Hz and 0.50 Hz, respectively.

2011 earthquake. The intensity measures for each simulation without considering the variation with distance also show a general reduction in values in the 3D simulations compared to 1D simulations.

This study also shows that the reduction of the PGD residuals in the 3D simulations is a combined effect of both shallow and deep 3D structures especially above certain hypocentral distances. Incorporating only the upper 30 km 3D structure will still improve the fit to the observed PGD values. Lastly, depending on the level of desired model accuracy and available computational resources, the 0.25 Hz SW4 3D simulations may be sufficient to model the kinematics and time-dependent crustal deformation measured by GNSS. Our results demonstrate that future studies of time-dependent crustal deformation should consider using 3D structure or Green's functions, in particular when peak intensity measures such as PGD are the most critical.

Acknowledgements

This work was funded by NASA ROSES grant 80NSSC21K0841. SW4 simulations were performed on the Lassen graphics processing unit accelerated platform operated by Livermore Computing using a Computing Grand Challenge allocation and on Talapas High Performance Computing at the University of Oregon. This work was performed in part under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. We thank the Computational Infrastructure for Geodynamics (http://geodynamics.org) which is funded by the National Science Foundation under awards EAR-0949446 and EAR-1550901. We also appreciate Sui Tung and an anonymous reviewer and editor Mathilde Radiguet for their constructive feedback and suggestions that helped improve the quality of this manuscript.

6 Data and Code Availability

SW4 is an open-source code and available at https://github.com/geodynamics/sw4 (last accessed January 2023) hosted by the Computational Infrastructure for Geodynamics (http://geodynamics.org). MudPy and FakeQuakes are available at https://github.com/oluwaseunfadugba/MudPy. Our maps were made with PyGMT (Uieda et al., 2021) available at https://github.com/GenericMappingTools/pygmt. PyGMT wraps around GMT6 (Wessel et al., 2019). We used the Slab2.0 model (Hayes, 2018) to create a fault geometry mesh for the Japan Trench using Gmsh (Geuzaine and Remacle, 2009). Our figures were

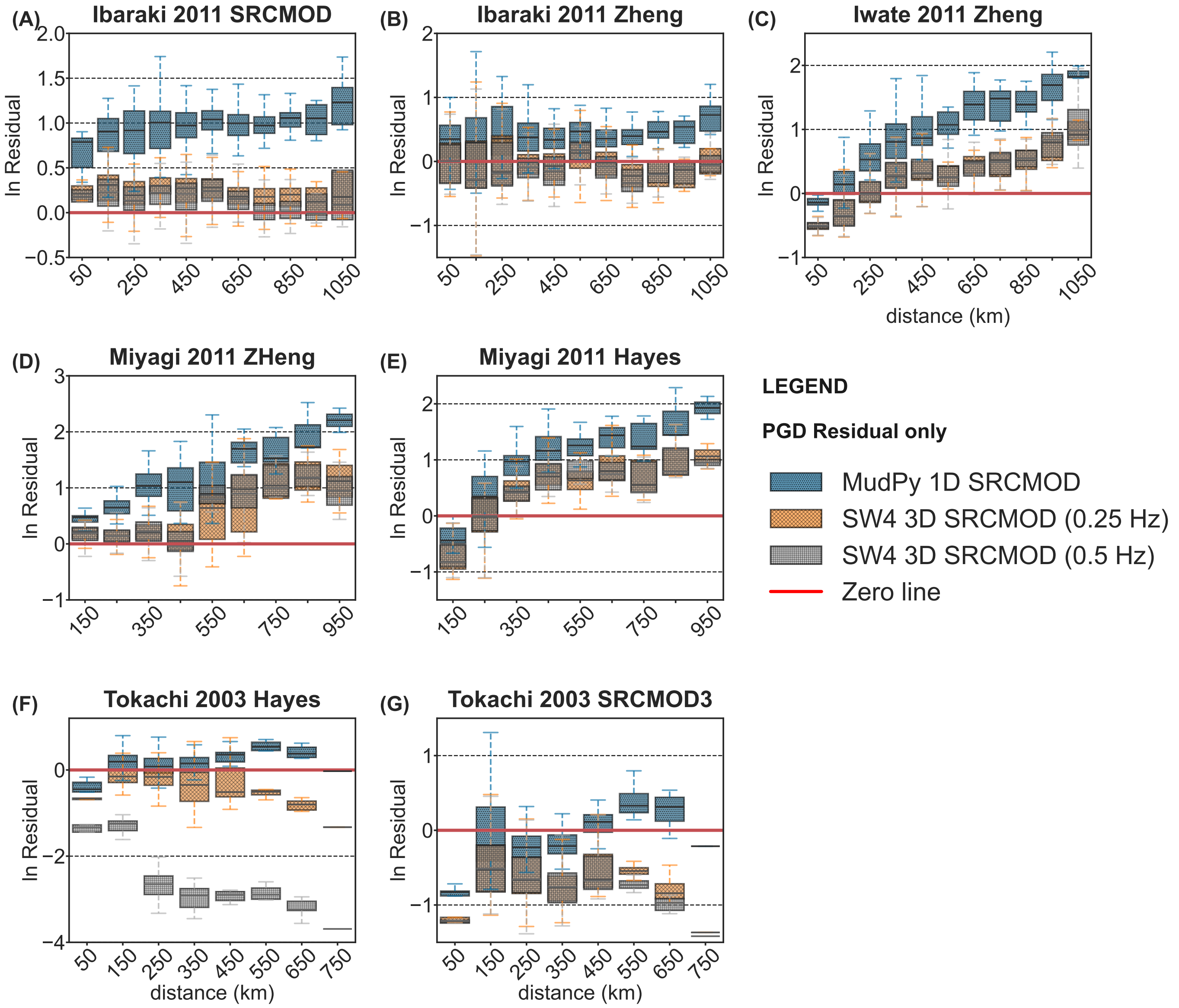


Figure 14 Effect of the maximum frequency (f_{max}) on PGD for all the simulations. The blue boxplots (circle hatched style) represent the MudPy 1D residuals. The orange boxplots (diamond hatch style) represent the SW4 3D simulation with fmax of 0.25 Hz. The gray boxplots (crossed hatched style) represent the SW4 3D simulation with fmax of 0.50 Hz. The red horizontal line represents the zero residual line.

made with Python 3 (Van Rossum and Drake, 2009), Seaborn (Waskom, 2021), Pandas (McKinney, 2010), and ObsPy (Beyreuther et al., 2010). Our codes are available at https://github.com/oluwaseunfadugba/1D_vs_3D_HR-GNSS_CrustalDeformation. We downloaded the JIVSM (Koketsu et al., 2008, 2009), which is the basis for our 3D modeling, from the Headquarters for Earthquake Research Promotion of Japanese Government (https://www.jishin.go.jp/evaluation/seismic_hazard_map/lpshm/) on 10/14/2021 in two overlapping sections: West Japan (https://www.jishin.go.jp/main/chousa/12_choshuki/dat/nankai/lp2012nankai-w_str.zip) and East Japan (https://www.jishin.go.jp/main/chousa/12_choshuki/dat/nankai/lp2012nankai-e_str.zip), each comprises 23 layers. The physical property values of the layers are from https://www.jishin.go.jp/main/chousa/12_choshuki/dat/nankai/lp2012nankai_str_val.pdf. The version provided here is not the original version published by JIVSM and is instead a modified version. The GNSS stations, mesh, 3D velocity model, projected rupture models for each earthquake on the Japan Trench mesh, the codes at the time of publication and the corresponding 100 realizations of the mean rupture models generated using FakeQuakes are available on Zenodo (Fadugba et al., 2023).

7 Competing Interests

The authors have no competing interests.

References

Aagaard, B. T., Knepley, M. G., and Williams, C. A. A domain decomposition approach to implementing fault slip in finite-element models of quasi-static and dynamic crustal deformation. *Journal of Geophysical Research: Solid Earth*, 118(6): 3059–3079, June 2013. doi: 10.1002/jgrb.50217.

Baltay, A. S., Hanks, T. C., and Abrahamson, N. A. Uncertainty, Variability, and Earthquake Physics in Ground-Motion Prediction Equations. *Bulletin of the Seismological Society of America*, May 2017. doi: 10.1785/0120160164.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., and Wassermann, J. ObsPy: A Python Toolbox for Seismology. *Seismological Research Letters*, 81(3):530–533, May 2010. doi: 10.1785/gssrl.81.3.530.

Blaser, L., Kruger, F., Ohrnberger, M., and Scherbaum, F. Scaling Relations of Earthquake Source Parameter Estimates with Special Focus on Subduction Environment. *Bulletin of the Seismological Society of America*, 100(6):2914–2926, Dec. 2010. doi: 10.1785/0120100111.

Bock, Y., Melgar, D., and Crowell, B. W. Real-Time Strong-Motion Broadband Displacements from Collocated GPS and Accelerometers. *Bulletin of the Seismological Society of America*, 101(6):2904–2925, Dec. 2011. doi: 10.1785/0120110007.

Delouis, B., Nocquet, J., and Vallée, M. Slip distribution of the February 27, 2010 Mw = 8.8 Maule Earthquake, central Chile, from static and high-rate GPS, InSAR, and broadband teleseismic data. *Geophysical Research Letters*, 37(17), Sept. 2010. doi: 10.1029/2010gl043899.

Dziewonski, A. M. and Anderson, D. L. Preliminary reference Earth model. *Physics of the Earth and Planetary Interiors*, 25(4): 297–356, June 1981. doi: 10.1016/0031-9201(81)90046-7.

Fadugba, O. I., Sahakian, V. J., Melgar, D. M., Rodgers, A., and Shimony, R. The Impact of the Three-Dimensional Structure of a Subduction Zone on Time-dependent Crustal Deformation Measured by HR-GNSS, 2023. doi: 10.5281/zenodo.7765170.

Furumura, T. and Kennett, B. L. N. On the nature of regional seismic phases-III. The influence of crustal heterogeneity on the wavefield for subduction earthquakes: the 1985 Michoacan and 1995 Copala, Guerrero, Mexico earthquakes. *Geophysical Journal International*, 135(3):1060–1084, Dec. 1998. doi: 10.1046/j.1365-246x.1998.00698.x.

Furumura, T. and Kennett, B. L. N. Subduction zone guided waves and the heterogeneity structure of the subducted plate: Intensity anomalies in northern Japan. *Journal of Geophysical Research: Solid Earth*, 110(B10), Oct. 2005. doi: 10.1029/2004jb003486.

Furumura, T. and Singh, S. Regional Wave Propagation from Mexican Subduction Zone Earthquakes: The Attenuation Functions for Interplate and Inslab Events. *Bulletin of the Seismological Society of America*, 92(6):2110–2125, Aug. 2002. doi: 10.1785/0120010278.

Geng, J., Pan, Y., Li, X., Guo, J., Liu, J., Chen, X., and Zhang, Y. Noise Characteristics of High-Rate Multi-GNSS for Subdaily Crustal Deformation Monitoring. *Journal of Geophysical Research: Solid Earth*, 123(2):1987–2002, Feb. 2018. doi: 10.1002/2018jb015527.

Geuzaine, C. and Remacle, J. Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79(11): 1309–1331, May 2009. doi: 10.1002/nme.2579.

Goldberg, D., Melgar, D., Hayes, G., Crowell, B., and Sahakian, V. A Ground-Motion Model for GNSS Peak Ground Displacement. *Bulletin of the Seismolgical Society of America*, 111, 2021. doi: 10.1785/0120210042.

Goldberg, D. E. and Melgar, D. Generation and Validation of Broadband Synthetic P Waves in Semistochastic Models of Large Earthquakes. *Bulletin of the Seismological Society of America*, 110(4):1982–1995, July 2020. doi: 10.1785/0120200049.

Goldberg, D. E., Koch, P., Melgar, D., Riquelme, S., and Yeck, W. L. Beyond the Teleseism: Introducing Regional Seismic and Geodetic Data into Routine USGS Finite-Fault Modeling. *Seismological Research Letters*, 93(6):3308–3323, Aug. 2022. doi: 10.1785/0220220047.

Graves, R. Standard Rupture Format Version 2.0, 2014.

Graves, R. and Pitarka, A. Refinements to the Graves and Pitarka (2010) Broadband Ground-Motion Simulation Method. *Seismological Research Letters*, 86(1):75–80, Dec. 2014. doi: 10.1785/0220140101.

Graves, R. W. and Pitarka, A. Broadband Ground-Motion Simulation Using a Hybrid Approach. *Bulletin of the Seismological Society of America*, 100(5A):2095–2123, Sept. 2010. doi: 10.1785/0120100057.

Hartzell, S., Harmsen, S., and Frankel, A. Effects of 3D Random Correlated Velocity Perturbations on Predicted Ground Motions. *Bulletin of the Seismological Society of America*, 100 (4):1415–1426, July 2010. doi: 10.1785/0120090060.

Hayes, G. Slab2 - A Comprehensive Subduction Zone Geometry Model, 2018. doi: 10.5066/F7PV6JNV.

Hayes, G. P. The finite, kinematic rupture properties of great-sized earthquakes since 1990. *Earth and Planetary Science Letters*, 468:94–100, June 2017. doi: 10.1016/j.epsl.2017.04.003.

Hearn, E. and Burgmann, R. The Effect of Elastic Layering on Inversions of GPS Data for Coseismic Slip and Resulting Stress Changes: Strike-Slip Earthquakes. *Bulletin of the Seismological Society of America*, 95(5):1637–1653, Oct. 2005. doi: 10.1785/0120040158.

Ide, S. *Slip Inversion*, page 193–223. Elsevier, 2007. doi: 10.1016/b978-044452748-6.00068-7.

Kaneko, Y., Ito, Y., Chow, B., Wallace, L. M., Tape, C., Grapenthin, R., D'Anastasio, E., Henrys, S., and Hino, R. Ultra-long Duration of Seismic Ground Motion Arising From a Thick, Low-Velocity Sedimentary Wedge. *Journal of Geophysical Research: Solid Earth*, 124(10):10347–10359, Oct. 2019. doi: 10.1029/2019jb017795.

Koketsu, K., Hikima, K., Miyazaki, S., and Ide, S. Joint inversion of strong motion and geodetic data for the source process of the 2003 Tokachi-oki, Hokkaido, earthquake. *Earth, Planets and Space*, 56(3):329–334, 2004. doi: 10.1186/bf03353060.

Koketsu, K., Miyake, H., Fujiwara, H., and Hashimoto, H. Progress towards a Japan integrated velocity structure model and long-period ground motion hazard map,. *Proceedings of the 14th World Conference on Earthquake Engineering, Paper*, 2008.

Koketsu, K., Miyake, H., Afnimar, and Tanaka, Y. A proposal for a standard procedure of modeling 3-D velocity structures and its application to the Tokyo metropolitan area, Japan. *Tectonophysics*, 472(1–4):290–300, July 2009. doi: 10.1016/j.tecto.2008.05.037.

Kolmogorov, A. Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Istituto Italiano degli Attuari*, 4:83–91 5, 1933.

Kotha, S. R., Weatherill, G., Bindi, D., and Cotton, F. A regionally-adaptable ground-motion model for shallow crustal earthquakes in Europe. *Bulletin of Earthquake Engineering*, 18(9): 4091–4125, May 2020. doi: 10.1007/s10518-020-00869-1.

Kubo, H., Asano, K., and Iwata, T. Source-rupture process of the 2011 Ibaraki-oki, Japan, earthquake (Mw 7.9) estimated from the joint inversion of strong-motion and GPS Data: Relationship with seamount and Philippine Sea Plate. *Geophysical Research Letters*, 40(12):3003–3007, June 2013. doi: 10.1002/grl.50558.

Kuehn, N. M. and Abrahamson, N. A. Spatial correlations of ground motion for non-ergodic seismic hazard analysis. *Earthquake Engineering & Structural Dynamics*, 49(1):4–23, Oct. 2019. doi: 10.1002/eqe.3221.

Landwehr, N., Kuehn, N. M., Scheffer, T., and Abrahamson, N. A Nonergodic Ground-Motion Model for California with

Spatially Varying Coefficients. *Bulletin of the Seismological Society of America*, 106(6):2574–2583, Oct. 2016. doi: 10.1785/0120160118.

Langer, L., Gharti, H. N., and Tromp, J. Impact of topography and three-dimensional heterogeneity on coseismic deformation. *Geophysical Journal International*, 217(2):866–878, Jan. 2019. doi: 10.1093/gji/ggz060.

Langer, L., Beller, S., Hirakawa, E., and Tromp, J. Impact of sedimentary basins on Green's functions for static slip inversion. *Geophysical Journal International*, 232(1):569–580, Aug. 2022. doi: 10.1093/gji/ggac344.

Laske, G., Masters, G., Ma, Z., and Pasyanos, M. Update on CRUST1.0 — A 1-degree global model of Earth's crust. *Geophysical research abstracts*, 15(15):2658, 2013.

LeVeque, R. J., Waagan, K., González, F. I., Rim, D., and Lin, G. Generating Random Earthquake Events for Probabilistic Tsunami Hazard Assessment. *Pure and Applied Geophysics*, 173(12): 3671–3692, Aug. 2016. doi: 10.1007/s00024-016-1357-1.

Maeda, T., Takemura, S., and Furumura, T. OpenSWPC: an open-source integrated parallel simulation code for modeling seismic wave propagation in 3D heterogeneous viscoelastic media. *Earth, Planets and Space*, 69(1), July 2017. doi: 10.1186/s40623-017-0687-2.

Mann, M. E. and Abers, G. A. First-Order Mantle Subduction-Zone Structure Effects on Ground Motion: The 2016 Mw 7.1 Iniskin and 2018 Mw 7.1 Anchorage Earthquakes. *Seismological Research Letters*, 91(1):85–93, Oct. 2019. doi: 10.1785/0220190197.

McKinney, W. Data Structures for Statistical Computing in Python. In *Proceedings of the 9th Python in Science Conference*, SciPy. SciPy, 2010. doi: 10.25080/majora-92bf1922-00a.

Melgar, D. and Bock, Y. Kinematic earthquake source inversion and tsunami runup prediction with regional geophysical data. *Journal of Geophysical Research: Solid Earth*, 120(5):3324–3349, May 2015. doi: 10.1002/2014jb011832.

Melgar, D. and Hayes, G. P. Characterizing large earthquakes before rupture is complete. *Science Advances*, 5(5), May 2019. doi: 10.1126/sciadv.aav2032.

Melgar, D., LeVeque, R. J., Dreger, D. S., and Allen, R. M. Kinematic rupture scenarios and synthetic displacement data: An example application to the Cascadia subduction zone. *Journal of Geophysical Research: Solid Earth*, 121(9):6658–6674, Sept. 2016. doi: 10.1002/2016jb013314.

Melgar, D., Crowell, B. W., Melbourne, T. I., Szeliga, W., Santillan, M., and Scrivner, C. Noise Characteristics of Operational Real-Time High-Rate GNSS Positions in a Large Aperture Network. *Journal of Geophysical Research: Solid Earth*, 125(7), June 2020a. doi: 10.1029/2019jb019197.

Melgar, D., Ganas, A., Taymaz, T., Valkaniotis, S., Crowell, B. W., Kapetanidis, V., Tsironi, V., Yolsal-Çevikbilen, S., and Öcalan, T. Rupture kinematics of 2020 January 24 Mw 6.7 Doğanyol-Sivrice, Turkey earthquake on the East Anatolian Fault Zone imaged by space geodesy. *Geophysical Journal International*, 223 (2):862–874, July 2020b. doi: 10.1093/gji/ggaa345.

Mena, B., Mai, P. M., Olsen, K. B., Purvance, M. D., and Brune, J. N. Hybrid Broadband Ground-Motion Simulation Using Scattering Green's Functions: Application to Large-Magnitude Events. *Bulletin of the Seismological Society of America*, 100(5A):2143–2162, Sept. 2010. doi: 10.1785/0120080318.

N.O.A.A. National Geophysical Data Center. ETOPO1 1 Arc-Minute Global Relief Model, 2009. Accessed [10/14/2021].

Olsen, K. B. Site Amplification in the Los Angeles Basin from Three-Dimensional Modeling of Ground Motion. *Bulletin of the Seismological Society of America*, 90(6B):S77–S94, Dec. 2000. doi: 10.1785/0120000506.

Ozawa, S., Nishimura, T., Suito, H., Kobayashi, T., Tobita, M., and Imakiire, T. Coseismic and postseismic slip of the 2011 magnitude-9 Tohoku-Oki earthquake. *Nature*, 475(7356): 373–376, June 2011. doi: 10.1038/nature10227.

Petersson, N. A. and Sjögreen, B. Stable and Efficient Modeling of Anelastic Attenuation in Seismic Wave Propagation. *Communications in Computational Physics*, 12(1):193–225, July 2012. doi: 10.4208/cicp.201010.090611a.

Petersson, N. A. and Sjögreen, B. Wave propagation in anisotropic elastic materials and curvilinear coordinates using a summation-by-parts finite difference method. *Journal of Computational Physics*, 299:820–841, Oct. 2015. doi: 10.1016/j.jcp.2015.07.023.

Petersson, N. A. and Sjögreen, B. geodynamics/sw4: SW4, version 2.01, 2017. doi: 10.5281/ZENODO.1063644.

Rodgers, A. J., Anders Petersson, N., Pitarka, A., McCallen, D. B., Sjogreen, B., and Abrahamson, N. Broadband (0–5 Hz) Fully Deterministic 3D Ground-Motion Simulations of a Magnitude 7.0 Hayward Fault Earthquake: Comparison with Empirical Ground-Motion Models and 3D Path and Site Effects from Source Normalized Intensities. *Seismological Research Letters*, 90(3): 1268–1284, Feb. 2019. doi: 10.1785/0220180261.

Rodgers, A. J., Pitarka, A., Pankajakshan, R., Sjögreen, B., and Petersson, N. A. Regional-Scale 3D Ground-Motion Simulations of Mw 7 Earthquakes on the Hayward Fault, Northern California Resolving Frequencies 0–10 Hz and Including Site-Response Corrections. *Bulletin of the Seismological Society of America*, 110 (6):2862–2881, Aug. 2020. doi: 10.1785/0120200147.

Ruhl, C. J., Melgar, D., Geng, J., Goldberg, D. E., Crowell, B. W., Allen, R. M., Bock, Y., Barrientos, S., Riquelme, S., Baez, J. C., Cabral-Cano, E., Pérez-Campos, X., Hill, E. M., Protti, M., Ganas, A., Ruiz, M., Mothes, P., Jarrín, P., Nocquet, J., Avouac, J., and D'Anastasio, E. A Global Database of Strong-Motion Displacement GNSS Recordings and an Example Application to PGD Scaling. *Seismological Research Letters*, 90(1):271–279, Oct. 2018. doi: 10.1785/0220180177.

Sahakian, V., Melgar, D., and Muzli, M. Weak Near-Field Behavior of a Tsunami Earthquake: Toward Real-Time Identification for Local Warning. *Geophysical Research Letters*, 46(16):9519–9528, Aug. 2019a. doi: 10.1029/2019gl083989.

Sahakian, V. J., Melgar, D., Quintanar, L., Ramírez-Guzmán, L., Pérez-Campos, X., and Baltay, A. Ground Motions from the 7 and 19 September 2017 Tehuantepec and Puebla-Morelos, Mexico, Earthquakes. *Bulletin of the Seismological Society of America*, Oct. 2018. doi: 10.1785/0120180108.

Sahakian, V. J., Baltay, A., Hanks, T. C., Buehler, J., Vernon, F. L., Kilb, D., and Abrahamson, N. A. Ground Motion Residuals, Path Effects, and Crustal Properties: A Pilot Study in Southern California. *Journal of Geophysical Research: Solid Earth*, 124(6): 5738–5753, June 2019b. doi: 10.1029/2018jb016796.

Sjögreen, B. and Petersson, N. A. A Fourth Order Accurate Finite Difference Scheme for the Elastic Wave Equation in Second Order Formulation. *Journal of Scientific Computing*, 52(1):17–48, Sept. 2011. doi: 10.1007/s10915-011-9531-1.

Smirnov, N. Table for Estimating the Goodness of Fit of Empirical Distributions. *The Annals of Mathematical Statistics*, 19(2): 279–281, June 1948. doi: 10.1214/aoms/1177730256.

Tung, S. and Masterlark, T. Sensitivities of Near-field Tsunami Forecasts to Megathrust Deformation Predictions. *Journal of Geophysical Research: Solid Earth*, 123(2):1711–1735, Feb. 2018. doi: 10.1002/2017jb015354.

Ueno, H., Hatakeyama, S., Aketagawa, T., Funasaki, J., and

Hamada, N. Improvement of hypocenter determination procedures in the Japan Meteorological Agency. *Quarterly Journal of Seismology*, 65:123 – 134, 2002.

Uieda, L., Tian, D., Leong, W. J., Toney, L., Schlitzer, W., Yao, J., Grund, M., Jones, M., Materna, K., Newton, T., Ziebarth, M., and Wessel, P. PyGMT: A Python interface for the Generic Mapping Tools, 2021. doi: 10.5281/ZENODO.4592991.

Van Rossum, G. and Drake, F. *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, 2009.

Vidale, J. and Helmberger, D. Elastic finite-difference modeling of the 1971 San Fernando, California earthquake. *Bulletin of the Seismological Society of America*, 78(1):122–141, 1988. doi: 10.1785/BSSA0780010122.

Wald, D. J. and Graves, R. W. Resolution analysis of finite fault source inversion using one- and three-dimensional Green's functions: 2. Combining seismic and geodetic data. *Journal of Geophysical Research: Solid Earth*, 106(B5):8767–8788, May 2001. doi: 10.1029/2000jb900435.

Wallace, L. M., Kaneko, Y., Hreinsdóttir, S., Hamling, I., Peng, Z., Bartlow, N., D'Anastasio, E., and Fry, B. Large-scale dynamic triggering of shallow slow slip enhanced by overlying sedimentary wedge. *Nature Geoscience*, 10(10):765–770, Sept. 2017. doi: 10.1038/ngeo3021.

Waskom, M. seaborn: statistical data visualization. *Journal of Open Source Software*, 6(60):3021, Apr. 2021. doi: 10.21105/joss.03021.

Wessel, P., Luis, J. F., Uieda, L., Scharroo, R., Wobbe, F., Smith, W. H. F., and Tian, D. The Generic Mapping Tools Version 6. *Geochemistry, Geophysics, Geosystems*, 20(11):5556–5564, Nov. 2019. doi: 10.1029/2019gc008515.

Williams, C. A. and Wallace, L. M. Effects of material property variations on slip estimates for subduction interface slow-slip events. *Geophysical Research Letters*, 42(4):1113–1121, Feb. 2015. doi: 10.1002/2014gl062505.

Wirth, E. A., Sahakian, V. J., Wallace, L. M., and Melnick, D. The occurrence and hazards of great subduction zone earthquakes. *Nature Reviews Earth & Environment*, 3(2):125–140, Jan. 2022. doi: 10.1038/s43017-021-00245-w.

Yagi, Y. Source rupture process of the 2003 Tokachi-oki earthquake determined by joint inversion of teleseismic body wave and strong ground motion data. *Earth, Planets and Space*, 56 (3):311–316, 2004. doi: 10.1186/bf03353057.

Yamanaka, Y. and Kikuchi, M. Source process of the recurrent Tokachi-oki earthquake on September 26, 2003, inferred from teleseismic body waves. *Earth, Planets and Space*, 55(12): e21–e24, 2003. doi: 10.1186/bf03352479.

Yue, H. and Lay, T. Inversion of high-rate (1 sps) GPS data for rupture process of the 11 March 2011 Tohoku earthquake (Mw 9.1). *Geophysical Research Letters*, 38(7), Apr. 2011. doi: 10.1029/2011gl048700.

Zhang, H., Melgar, D., Sahakian, V., Searcy, J., and Lin, J.-T. Learning source, path and site effects: CNN-based on-site intensity prediction for earthquake early warning. *Geophysical Journal International*, 231(3):2186–2204, Aug. 2022. doi: 10.1093/gji/ggac325.

Zheng, X., Zhang, Y., Wang, R., Zhao, L., Li, W., and Huang, Q. Automatic Inversions of Strong-Motion Records for Finite-Fault Models of Significant Earthquakes in and Around Japan. *Journal of Geophysical Research: Solid Earth*, 125(9), Sept. 2020. doi: 10.1029/2020jb019992.

Zhu, L. and Rivera, L. A. A note on the dynamic and static displacements from a point source in multilayered media. *Geophysical Journal International*, 148(3):619–627, Mar. 2002. doi: 10.1046/j.1365-246x.2002.01610.x.

doi:10.26443/seismica.v3i1.692

SEISMICA

Virtual Shake Robot: Simulating Dynamics of Precariously Balanced Rocks for Overturning and Large-displacement Processes

Zhiang Chen * [1], **Ramón Arrowsmith** [1], **Jnaneshwar Das** [1], **Christine Wittich** [2], **Christopher Madugo** [3], **Albert Kottke** [3]

[1]School of Earth and Space Exploration, Arizona State University, Tempe, USA, [2]Department of Civil and Environmental Engineering, University of Nebraska-Lincoln, Lincoln, USA, [3]Pacific Gas and Electric Company, San Francisco, USA

Author contributions: *Conceptualization*: Z.C., R.A., and J.D.. *Methodology*: Z.C., R.A., C.W.. *Software*: Z.C. and J.D.. *Investigation*: Z.C. and C.W.. *Writing - original draft*: Z.C.. *Writing - Review & Editing*: Z.C., R.A., C.W., and A.K.. *Funding acquisition*: R.A., C.M., A.K., and J.D..

Abstract Understanding the dynamics of precariously balanced rocks (PBRs) is important for seismic hazard analysis and rockfall prediction. Utilizing a physics engine and robotic tools, we develop a virtual shake robot (VSR) to simulate the dynamics of PBRs during overturning and large-displacement processes. We present the background of physics engines and technical details of the VSR, including software architecture, mechanical structure, control system, and implementation procedures. Validation experiments show the median fragility contour from VSR simulation is within the 95% prediction intervals from previous physical experiments, when PGV/PGA is greater than 0.08 s. Using a physical mini shake robot, we validate the qualitative consistency of fragility anisotropy between the VSR and physical experiments. By overturning cuboids on flat terrain, the VSR reveals the relationship between fragility and geometric dimensions (e.g., aspect and scaling ratios). The ground motion orientation and lateral pedestal support affect PBR fragility. Large-displacement experiments estimate rock trajectories for different ground motions, which is useful for understanding the fate of toppled PBRs. Ground motions positively correlate with large displacement statistics such as mean trajectory length, mean largest velocity, and mean terminal distance. The overturning and large displacement processes of PBRs provide complementary methods of ground motion estimation.

Non-technical summary Fragile geological features such as precariously balanced rocks (PBRs) may provide ground motion constraints for seismic hazard analysis. PBRs, many discovered close to infrastructure, may form rockfall hazards. Modeling PBR trajectory also helps understand the fate of toppled PBRs and the processes of rocky slope development. The dynamics of PBRs, however, are nonlinear and present many challenges to analyze. Utilizing robotic tools, we develop a virtual shake robot (VSR) in simulation to study the dynamics of PBRs during overturning and large-displacement processes with realistic material properties and terrains. Using the VSR, we demonstrate that the PBR fragility is affected by ground motion directions and lateral supporting pedestals, which have seldom been considered in previous studies. The results of large-displacement experiments indicate that increasing ground motions result in greater PBR transportation. Additionally, the VSR has the advantage of rapid deployment, which plays an important role in our rock detection-mapping-analysis paradigm that aims to automate rock mapping and analysis by leveraging robotics and machine learning technologies.

Production Editor:
Gareth Funning
Handling Editor:
Pablo Heresi
Copy & Layout Editor:
Hannah F. Mark

Received:
March 31, 2023
Accepted:
December 24, 2023
Published:
January 15, 2024

1 Introduction

Precariously balanced rocks (PBRs) are boulders balanced on and not fixed to a sub-horizontal pedestal. The balance configuration and contact physics define PBR fragility—probability for overturning by a stimulus, usually earthquake ground motions. Seismologists have studied the overturning responses of PBRs from ground motions for seismic hazard analysis (Housner, 1963; Brune, 1996; Shi et al., 1996; Anooshehpoor et al., 2004; Rood et al., 2020, 2022). PBR fragility provides an upper bound on the strength of the ground motions in the time interval since the PBRs became fragile (Brune et al., 2006; Anderson et al., 2014). In southern California, most PBRs have been fragile for thousands of years or longer (Brune et al., 2006; Rood et al., 2022). Studying PBRs allows ground motion estimation with long return times, which are much longer than the modern instrumental earthquake catalogs. Such long-history ground motion estimation is important for assessing hazards for critical facilities such as large dams, nuclear power plants, and nuclear waste repositories (Rood et al., 2020). In principle, PBRs allow a partial test of hazard curves obtained from other information, including geological appraisals of earthquakes and nearby faults (Rood et al., 2020). Hazard curves, outputs of probabilistic seismic hazard analysis (PSHA), ex-

*Corresponding author: zchen256@asu.edu

press the rate at which ground motions are equaled or exceeded as a function of the amplitude of the motion. PBRs constrain the hazard curves at very long return times.

Seismic hazard analysis typically considers the PBR overturning responses, which are immediate binary results (balanced or overturned, Anderson et al., 2014). However, their motions after overturning are complex and informative. Overturned PBRs can slide, rotate, rock, and bounce. The large displacements of these rocks contribute to understanding the fate of PBRs and the development of rocky slopes. Additionally, many discovered PBRs present rockfall hazards (Anderson et al., 2014). As a serious natural hazard, rockfall poses a major threat to infrastructure, transportation lines, and people (Dorren, 2003; Leine et al., 2013). Predicting rockfall trajectories in complex terrains is essential for implementing protective measures.

Using PBRs for seismic hazard analysis presents challenges in PBR mapping, PBR dynamics, PBR dating, and hazard modeling (Fig. 1). PBRs are not everywhere. PBR mapping locates and obtains PBR geometries and contact properties. Hundreds of PBRs have been manually located in southern California, and their metadata is archived at Southern California Earthquake Center (SCEC researchers, 2022). However, many of them were discovered near developed roads because of accessibility. The heterogeneous distribution indicates a sampling bias for ground motion estimations. A recent study developed unpiloted aerial vehicles (UAV) and onboard machine learning to search for PBRs (Chen et al., 2024, in prep). The PBR geometry is a critical factor affecting PBR dynamics. PBR geometry is often modeled by minimal contact angles from a 2D side-view photo, where a contact angle is an angle between the gravity vector and the connection from the mass centroid to a rocking point (Haddad et al., 2012; Shi et al., 1996). 3D PBR models are reconstructed using terrestrial laser scanning, structure from motion (SfM), or robotic real-time mapping technologies (Veeraraghavan et al., 2016; Rood et al., 2020; Chen et al., 2024, in prep). PBR dynamics focus on the response of PBRs from known ground motions (a forward dynamics problem). PBR forward dynamics are nonlinear and have been studied for over a century (Milne and Omori, 1893; Housner, 1963). The methods of PBR dating include cosmogenic, rock varnish, and quantitative geomorphic models (Bell et al., 1998; Rood et al., 2020). With the forward dynamics and ages of PBRs, hazard modeling integrates such information into a PSHA to test or rectify hazard curves. Despite the importance of all four challenges in PBR usage for hazard analysis, this paper specifically concentrates on modeling PBR dynamics through simulation tool development and experiments.

The dynamics of PBRs for seismic hazard analysis are aimed to model an overturning process, which is a mapping from ground motions to PBR response of overturned or not. Ground motions are characterized by intensity measures (Anderson et al., 2014; Purvance et al., 2008), such as peak ground acceleration (PGA), peak ground velocity (PGV), peak ground displacement, SA(1Hz), and SA(2 Hz), where SA(n Hz)

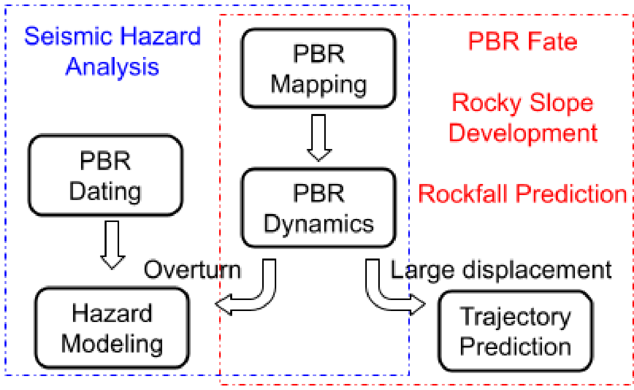


Figure 1 Workflow of using precariously balanced rocks (PBRs) for seismic hazard analysis, PBR fate, rocky slope development, and rockfall prediction. PBR dynamics involve overturning process and large-displacement process.

represents the peak acceleration response of a single-degree-of-freedom oscillator with undamped natural frequency of n Hz and 5% damping to the ground motions (Baker et al., 2021). Purvance et al. (2008) found that PGV/PGA and PGA were the strongest indicators of the overturning potentials among the above intensity measures. Since then, PGV/PGA and PGV have been commonly used as ground motion inputs to study the overturning problem (e.g., Veeraraghavan et al., 2016; Rood et al., 2020). The PBR overturning response can be described as a function of PGV/PGA and PGA,

$$OR = f(PGV/PGA, PGA) \tag{1}$$

where OR is a binary variable with value 1 indicating overturned response and value 0 indicating balanced response. Note that in the real world the PBR dynamics are deterministic, which means, given a ground motion, a PBR response can only be either overturned or balanced. In this case, if we uniformly discretize $(PGV/PGA, PGA)$ ground motion space and assume each ground motion has the same probability, a PBR is *more fragile* than another when it has more overturned responses. To consider uncertainties in PBR dynamics and to integrate the PBR dynamics model into PSHA, a probabilistic model that indicates the probability of being overturned can be obtained from Monte Carlo simulation and logistic regression (Purvance, 2005).

The overturning dynamics of PBRs have been studied for over a century. Assuming sufficiently large friction, previous studies modeled 2D rotation motion of rectangles (vertical plane) (Milne and Omori, 1893; Kirkpatrick, 1927; Housner, 1963; Yim et al., 1980; Purvance et al., 2008). Other studies explored the models and criteria for more complicated motions such as rotating, sliding, bouncing, and rocking in 2D (Ishiyama, 1982; Scalia and Sumbatyan, 1996; Shenton and Harry, 1996; Pompei et al., 1998). Purvance et al. (2012) used the discrete element method (DEM) to analyze the overturning response of 3D PBR models. This method required many repeated experiments to calibrate parameters of stiffness and damping (Purvance et al., 2012; Saifullah and Wittich, 2022, 2021), and the simulation for each experiment is computationally expensive. Veeraraghavan et al. (2016) presented an alternative method to analyze the 3D PBR overturning dynamics using a constraint-based model. The constraint-based model formed a linear complementarity problem from contact constraints and solved contact impulses by an iterative numerical algorithm (Chapter 2 in Veeraraghavan, 2015). From the contact impulses, contact forces were computed to update object velocity. Their study (Veeraraghavan et al., 2016) validated the constraint-based model in agreement with the physical shake table experiments from Purvance et al. (2008). The constraint-based model from Veeraraghavan et al. (2016) is similar to the collision response stage in physics engines. However, physics engines directly apply the contact impulse to change object velocity instead of computing intermediate contact forces (see Section 2.1). Modern physics engines are also enhanced with efficient algorithms and Graphic Processing Unit (GPU) hardware accelerations.

Besides the overturning dynamics, we also investigate large-displacement dynamics of PBRs, which are important for PBR fate study, rocky slope development understanding, and rockfall prediction. The goal of the large-displacement dynamics is to predict trajectories of PBRs after being overturned. PBR trajectories are affected by factors including PBR initial state, PBR physics properties, terrain morphology, and terrain physics properties. Given the same configurations for all the other factors, a PBR trajectory is distinguished from its initial state,

$$T = h(s) \tag{2}$$

where T is the trajectory (position and orientation with time), h is a function that maps an initial state to a trajectory, and s is the initial state such as position, orientation, and initial velocity. Compared with the overturning dynamics that have been widely explored in previous PBR dynamics studies, general large-displacement dynamics of rocks were more studied in rockfall hazard applications. Early studies restricted rockfall motions to 2D vertical planes and built mathematical models to describe discrete motion modes (Bozzolo and Pamini, 1986; Kobayashi et al., 1990; Azzoni et al., 1995). 3D rockfall models were developed to simulate particle interactions with digital elevation models and digital terrain models (Gascuel et al., 1999; Agliardi and Crosta, 2003; Lan et al., 2007; Guzzetti et al., 2002; Caviezel et al., 2019). Recently, Hao et al. (2021) used a physics engine to simulate rockfall trajectories on a terrain model that was reconstructed by aerial photographs from unpiloted aircraft systems (UAS).

By integrating advanced technologies from physics engines and robotics, we have developed a virtual shake robot (VSR) to facilitate the study of the overturning and large-displacement dynamics of PBRs. Our VSR relies on three core technologies: Robot Operating System (ROS), Gazebo simulation toolbox, and Bullet physics engine. ROS is a software platform that provides a set of libraries and tools for robot control and perception (Stanford Artificial Intelligence Laboratory, 2018). Gazebo is a simulation toolbox that provides a simple way to build a virtual world including virtual robots and environments (Koenig and Howard). The dynamics of the virtual world are managed by a physics engine. Gazebo supports four high-performance physics engines: Open Dynamics Engine, Bullet, Simbody, and DART. Because the Bullet physics engine has shown reliable simulation results in many scientific studies (Zhu and Zhao, 2019; Ma et al., 2018; Sun et al., 2019), we select Bullet as the physics engine for the VSR.

This study is aimed to advance the simulation of PBR dynamics by seamlessly addressing both overturning and large-displacement processes. Simulation tools are important for science studies, but technical nuances may affect experiment results. For example, rock behaviors such as rotating, sliding, jumping, and rocking affect the dynamics of a shake table. However, such details are unclear in previous studies (Purvance et al., 2012; Veeraraghavan et al., 2016). To reduce the effects of the coupling dynamics, the VSR implements a hierarchical control system.

Previously, simulation models for the overturning and large-displacement dynamics were built independently. Our VSR is the first tool for both dynamics studies. Additionally, complex terrain—anything other than a flat pedestal—can either increase or decrease the fragility of a PBR, depending on the specific characteristics of the surrounding terrain and the contact. Our VSR supports arbitrarily complex terrains, e.g., mesh models from UAS and SfM, to advance PBR dynamics studies. From shake simulations, we demonstrate that surrounding pedestals are critical to reduce PBR fragility. Lastly, we expect to see more robotic applications to PBR mapping, and thus the VSR can be integrated into autonomous mapping systems where ROS has widely been used.

To enhance the clarity of the article structure, we provide an outline as follows. Following the introduction in Section 1, we present technical advancements and review related work in Section 2. Because a physics engine is a new approach to simulating the PBR dynamics, we begin by reviewing the technical details in the Bullet physics engine and compare it with the DEM to establish the necessary context for our research. We also introduce relevant applications of physics engines. Moving on to Section 3, we provide a comprehensive description of the VSR system, covering its software architecture, mechanical structure, control system, and implementation procedures. For validation purposes, Section 4 compares the simulation results from the VSR with physical experiment data from a previous study, as well as with data collected from a physical, mini shake robot. In Section 5, we present our simulation experiments, with a focus on scientific applications related to the overturning and large-displacement processes. In Section 6, we discuss the validation experiments, the overturning and large-displacement experiments, as well as the limitations and prospects for future work. Finally, we summarize this study in Section 7. Throughout this paper, we use the terminology 'robot' to refer to both the VSR and mini shake robot because of their reliance on robotic concepts, including control and perception modules, as well as the use of robotic tools such as ROS. It is worth emphasizing that both robots facilitate the automation of data collection, making them valuable tools for obtaining data in earthquake studies.

2 Background

2.1 Bullet Physics Engine

The Bullet physics engine simulates rigid body dynamics by computing rigid body states (poses and velocities) in a discrete simulation loop with a fixed small time step in the range of 30 Hz to 10 kHz (Coumans and Bai, 2016).

The Bullet rigid body simulation loop includes four stages, as shown in Fig. 2a. The collision detection stage predicts contact points (where to contact) and time of impact (when to contact) using efficient algorithms (van den Bergen, 2003; Williams et al., 2014). When collision is predicted within a simulation time step, the collision response stage computes the impulse from the collision. The forward dynamics stage computes external force, torque, and inertia. Finally, the numerical time integration stage updates position and linear velocity of each object using semi-explicit Euler integration method,

$$v_{t+\Delta t} = v_t + a\Delta t = v_t + \frac{F_{ext} + F_c}{m}\Delta t, \quad (3)$$

$$x_{t+\Delta t} = x_t + v_{t+\Delta t}\Delta t \quad (4)$$

where v $[m \cdot s^{-1}]$ is velocity, t $[s]$ is the current time, Δt $[s]$ is the simulation time step, a $[m \cdot s^{-2}]$ is acceleration, F_{ext} $[kg \cdot m \cdot s^{-2}]$ is total external force, F_c $[kg \cdot m \cdot s^{-2}]$ is total contact force, m $[kg]$ is object mass, and x $[m]$ is object position. F_{ext} can be gravity, wind force field, and user-defined force, which are computed at the forward dynamics stage. F_c includes collision force and friction, which have non-trivial solutions. Instead of integrating contact force over time in Eq. 3, Bullet calculates the total impulse to update velocity,

$$F_c\Delta t = J, \quad (5)$$

$$v_{t+\Delta t} = v_t + \frac{F_{ext}}{m}\Delta t + \frac{J}{m}. \quad (6)$$

where J $[kg \cdot m \cdot s^{-1}]$ is the total impulse from the contact, which is computed at the collision response stage. Similarly to the linear equations, the angular equations update angular velocity and orientation by considering torque, inertia, and angular acceleration. Such a method of computing contact impulse was known as impulse-based dynamics (IBD), introduced by Mirtich and Canny (1995) and improved by Bender et al. (2013). Bullet computes the contact impulse by modeling collision dynamics as equality and inequality constraints to form a mixed linear complementarity problem (MLCP). The constraints from a collision include contact constraints, friction constraints, and joint constraints. The projected Gauss-Seidel algorithm (Erleben et al., 2005) solves the MLCP by iteratively approximating an impulse until all the constraints are satisfied (converged) or a maximum number of iterations is met.

2.2 Physics Engine versus Discrete Element Method (DEM)

Classical DEM discretizes each object as a collection of small spheres (Cundall and Strack, 1979). For each fixed small time step in discrete simulation loops, as illustrated in Fig. 2b, DEM computes the states of all spherical particles. The rationale of DEM is that the time step chosen is so small that during a single time step disturbances cannot propagate from any spherical particles further than its immediate neighbors. The total force on each spherical particle is only determined by external force defined by the user and contact force imparted by its neighbors with which it is in contact. DEM allows penetration (geometric overlaps) between spherical particles, and thus does not need collision detection to predict when and where to collide. When spherical particles contact, DEM uses a force-displacement model to compute contact force from penetration depth. The processes of the forward dynamics stage and numerical time integration stage in DEM are similar to the processes in Bullet. While originally intended to represent granular media with spherical particles, DEM has since been extended to simulate the behavior of arbitrary rigid bodies (meshes) using polyhedral particles (Cundall, 1988; Itasca Consulting Group, Inc., 2020). However, DEM computes the contact forces at all particles' penetrating faces and nodes, which is different from Bullet where the collision computation is based on each individual mesh object. In DEM, a more highly discretized surface within an object provides a more accurate representation of the contact force distribution. However, increasing the number of discretized particles significantly increases the simulation time.

Although the formulation of physics engines and DEM were motivated by different purposes, their modern applications involve various fields of engineering and science. Physics engines were originally developed to rapidly simulate physical processes in computer games and animations (Millington, 2007). With an increase in accuracy, physics engines quickly were used in engineering and scientific studies including robotics (Drumwright et al., 2010; Erez et al., 2015), agricultural machinery safety (Sun et al., 2019), construction materials (Garcia-Hernandez et al., 2021), earthquake studies (Xu et al., 2013; Kim et al., 2015), rockfall hazard zoning (Hao et al., 2021), and granular soil studies (He et al., 2021; Izadi and Bezuijen, 2014; Pytlos et al., 2015; Toson and Khinast, 2017). Whereas physics engine applications involve both individual objects and particle assemblages, DEM has primarily focused on modeling mechanical behavior of particle assemblages since it was introduced (Cundall and Strack, 1979). Modern DEM applications include particle transportation in mining (Pezo et al., 2015), particle compaction in material science (Martin et al., 2003), fertilizer spreading in agriculture (Coetzee and Lombard, 2011), soil-tool interaction (Asaf et al., 2007; Catanoso et al., 2020), particle mixer and grinder machinery (Alian et al., 2015; Cleary, 2015), rocky slopes in geoengineering (Chuhan et al., 1997), dynamic analysis of infrastructures (Çaktı et al., 2016; Mehrotra and DeJong, 2017; Papantonopoulos et al., 2002), and fault ruptures (Garcia and Bray, 2018, 2022; Chiama et al., 2023).

The difference in the contact models of the Bullet physics engine and DEM is a major factor affecting their

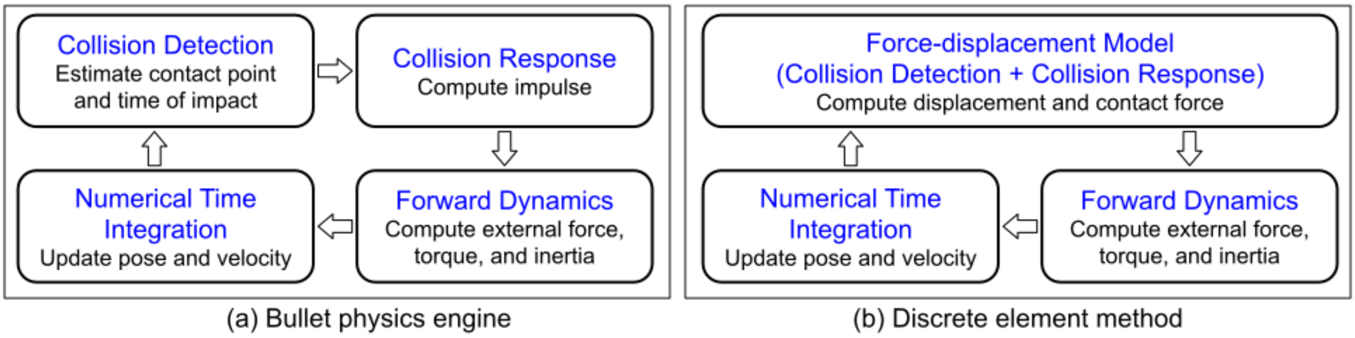


Figure 2 Simulations loops of (a) Bullet physics engine and (b) discrete element method (DEM). In both Bullet physics engine and DEM, states of objects are updated in each sequential discrete simulation loop with a fixed small time step. Their mechanisms to process collision are fundamentally different.

computational efficiency. Bullet applies a hard contact model where penetrations are not allowed between any two rigid bodies. Thus, in its simulation loop, Bullet needs the collision detection stage to predict when and where to contact, and then collision impulses are computed in the collision response stage. The hard contact model is formalized as a MLCP such that iterative numerical methods (e.g., projected Gauss-Seidel algorithm) are leveraged to efficiently compute collision impulses. The Bullet physics engine requires users to provide macro physics parameters, including restitution, Coulomb friction, and rolling friction coefficients (Chen, 2022). On the other hand, DEM adopts a soft contact model (force-displacement model) where penetrations are allowed between two directly contacting particles. The soft contact model requires user-defined parameters such as stiffness, damping constant, and friction coefficient. To simulate rigid bodies, the stiffness is usually set at a very large value to reduce macroscopic deformation. Because of the high stiffness, the simulation time step must be small to yield small elastic rebound at each iteration to ensure numerical stability, which significantly increases the simulation time. Additionally, the hard contact model treats each object as an individual entity (e.g., using polygon mesh model), whereas DEM discretizes each object into small particles. The computational time and memory of DEM increase markedly with the number of particles. With the same computational hardware and settings, previous studies found physics engines were 10-250 times faster than DEM to achieve similar results (He et al., 2020, 2021). Given the fact that the soft contact model in DEM is a hypothetical model, the user-defined parameters within this model lack direct connections to macro physics properties. These parameters can be calibrated through an iterative process of adjusting parameter values and matching experimental observations. In contrast, the parameters required in Bullet represent macrophysics properties and may be directly measured through experiments.

Physics engines and DEM are numerical methods, neither of which is a true representation of reality, and both of which need calibration. Our goal here is to promote physics engine applications in PBR dynamics studies, which provides one more option with some advantages relative to DEM. More research is needed to compare the performance of physics engines and DEM on this topic.

2.3 Physics Engine Applications

In this subsection, we review physics engine applications related to this study. Physics engines have succeeded in simulating behavior of granular assemblies. Izadi and Bezuijen (2014) used Bullet to simulate the behavior of granular materials subjected to pluviation and vibration. Their simulation results were within the range of repeated laboratory experiments. Toson and Khinast (2017) applied Bullet to study quasi-static granular flows of non-spherical particles. While their Bullet simulation results for spherical particles were in agreement with DEM simulation results, implementing non-spherical particle simulation in Bullet was easier because simulations of non-spherical particles in DEM required an advanced discretization method (e.g., Lu et al., 2015). Their Bullet simulation results of non-spherical particles agreed with the prediction from the empirical Beverloo equation (Beverloo et al., 1961). He et al. (2020) compared physics engine and DEM simulations in granular soil behavior and showed that the physics engine achieved similar results to those of the DEM in a significantly shorter time. Komaragiri et al. (2021) demonstrated that the compaction behavior of asphalt mixture from Bullet simulation was very similar to the compaction behavior recorded from laboratory measurements.

Zhu and Zhao (2019) demonstrated the benefits of utilizing physics engines in material analysis. Their study employed the Bullet physics engine and integrated peridynamics to simulate crushable granular materials under mechanical loading. It was difficult to analyze particle breakage using DEM because traditional spherical particles were incapable of approximating particles with sharp corners and edges (Zhu and Zhao, 2019). The Bullet simulation results were consistent with experimental observations on normal compression line, particle size distribution, fractal dimension, and particle morphology.

The Bullet physics engine demonstrated reliable capability of simulating overturning and large-displacement dynamics (Ma et al., 2018; Sun et al., 2019; Hao et al., 2021). Ma et al. (2018) applied Bullet to simulate rocking dynamics of cuboids with various sizes and aspect ratios (height-to-width ratios). Their study demonstrated that the response of rocking blocks in Bullet simulation was consistent with analytical solutions of Housner equations (Housner, 1963). Sun et al. (2019) used Bullet to analyze overturning dynamics of agricultural tractors on a bank slope and on a uniform slope. The Bullet simulation results were similar to previous reports and also in reasonably good agreement with experimental results. Hao et al. (2021) simulated rockfall trajectories on terrains using PhysX physics engine. They showed the advantage of using the physics engine to simulate rock interactions with a high-resolution, realistic terrain model that was reconstructed by UAV aerial photographs.

3 Virtual Shake Robot

As shown in Fig. 3, we developed the VSR using ROS, Gazebo simulation toolbox, and Bullet physics engine. Utilizing various libraries and tools available in the ROS ecosystem, we built robot software composed of control and perception modules. The control module computed actuation forces for the VSR; the perception module monitored PBR states and ground motion. Using Gazebo, we designed the mechanical structure of the VSR and defined the general physics properties, including gravity and lighting. Gazebo simulation toolbox supports four physics engine options to manage the dynamics of the virtual world. We selected the Bullet physics engine for its reliable performance in many previous scientific studies (see Section 2.3; Zhu and Zhao, 2019; Ma et al., 2018; Sun et al., 2019). Implementing such a software architecture provides two main advantages. First, rather than directly interacting with the details in a physics engine, Gazebo only requires XML format configuration files where the user can select the desired physics engine and configure physics parameters. Once the configuration files are properly set, Gazebo passes the parameters to the physics engine, simplifying the user's experience and making it easy to switch between physics engines if necessary. Second, as illustrated in Fig. 3, leveraging ROS in the development of the robot software allowed us to reuse the software for both virtual and physical shake robots (Chen et al., 2022). The same robot software ensures that ground motion generation processes are consistent, which is critical to compare simulation and physical experiments.

The VSR has a straightforward mechanical structure, as illustrated in Fig. 4, composed of a base, linear rail, and pedestal. All three components are rigidly fixed, with the base anchored in the virtual world and the linear rail attached to the base. A prismatic joint, functioning as a prismatic motor, links the linear rail and pedestal and generates translational force to actuate the pedestal. This translational force is calculated from the control module. The VSR enables one-dimensional prismatic, horizontal ground motions under the constraint of the prismatic joint (Fig. 4). The VSR supports various pedestal models, including a flat terrain (Fig. 4a) and realistic terrains that were mapped from UAS and SfM (Fig. 4b, c). Switching a pedestal model is as simple as configuring the model path in a Gazebo configuration file.

We developed a hierarchical control module for the VSR (Fig. 5a). The control module provides two ground motion options. The first one is a single-pulse cosine ground displacement,

$$d(t) = A - Acos(2\pi ft), \tag{7}$$

where $d(t)$ is the ground displacement function, A is the amplitude, f is the frequency, and $t \in [0, 1/f]$ is time. A and f are derived from PGV/PGA and PGA,

$$f = \frac{1}{2\pi r}, \tag{8}$$

$$A = \frac{ag}{4\pi^2 f^2}. \tag{9}$$

where r is PGV/PGA, a is PGA, and g is the gravitational acceleration. As shown in Fig. 5a, the motion interpreter converts PGV and PGA to f and A using Eq. 8 and Eq. 9. Based on A and f, the trajectory planner takes the derivative of the cosine displacement function (Eq. 7) to obtain ground velocity function,

$$v(t) = \dot{d}(t) = 2\pi Afsin(2\pi ft), \tag{10}$$

where $v(t)$ is the ground velocity function, and $t \in [0, 1/f]$ is time. We uniformly discretized $v(t)$ to sample a set of velocities, $\{v\}$, as the input of the velocity controller. The sampling frequency is a user-defined parameter, usually between 100 and 200 Hz.

In addition to the cosine ground displacement function, the VSR supports ground motion simulation from real seismometer records. As shown in Fig. 5a, a low-pass filter first removes the high-frequency noise in the raw acceleration data. The numerical integration produces velocities from the accelerations. Then a high-pass filter removes the low-frequency noise in the velocities, because the low-frequency velocity noise may accumulate displacement errors. Note that the output of the high-pass filter is a set of velocities $\{v\}$, which has the same format as the output from the trajectory planner. We utilize a shared velocity controller to process the desired velocity commands, simplifying the control software.

We implemented a PID velocity controller to generate force commands for the prismatic joint. The velocity commands for the PID controller are derived from either a cosine ground placement function or a seismometer record. The output force from the PID velocity controller actuates the pedestal. At the same time, Gazebo measures the actual velocity of the pedestal and feeds it back to the PID velocity controller. The actual velocity (velocity measurements) of the pedestal may be different from the desired velocity (velocity commands), resulting in velocity error,

$$e = v_d - v_m, \tag{11}$$

where v_m is the actual velocity from measurement, and v_d is the desired velocity from the higher level of the control module. One reason for the velocity error is that PBR overturning behaviors may produce collision impacts that affect the pedestal dynamics. The objective

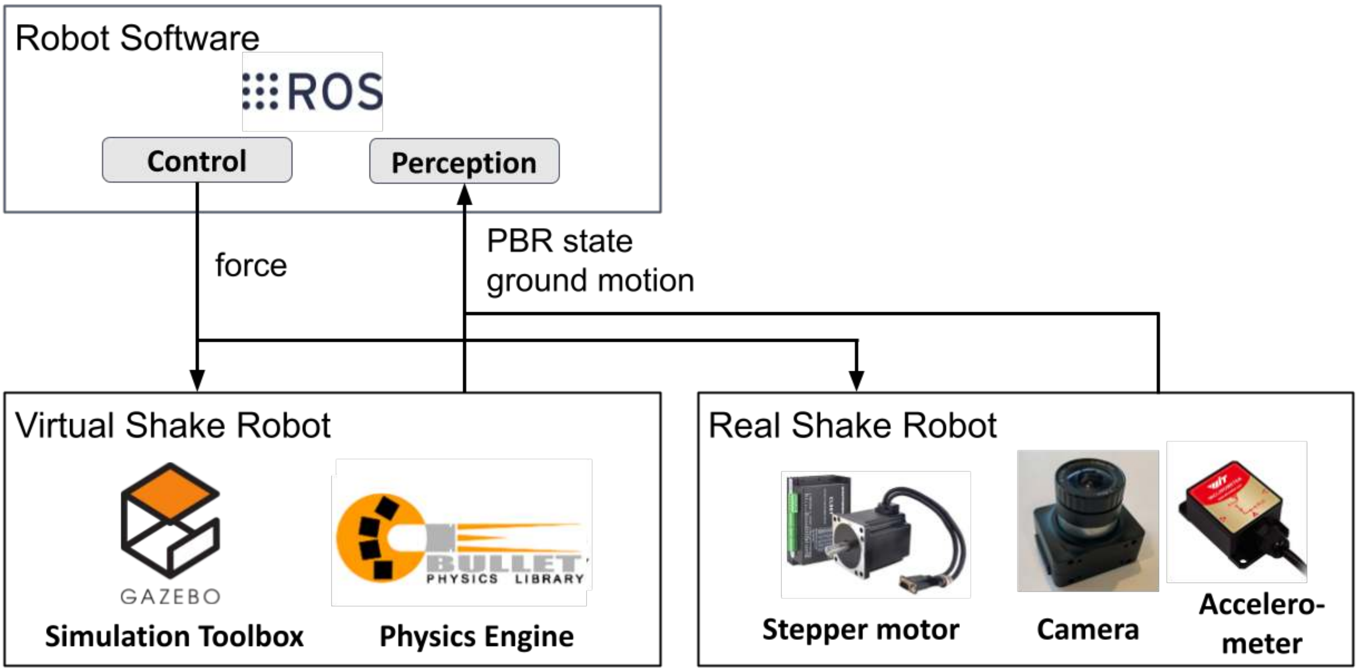


Figure 3 Software architecture of the virtual shake robot (VSR), composed of Robot Operating System (ROS), Gazebo simulation toolbox, and Bullet physics engine. Developing robot software based on ROS allows the reuse of robot software for both virtual and real shake robots, ensuring the ground motion generation processes remain consistent between the two environments.

of the PID velocity controller is to generate force for the prismatic joint to minimize the velocity error,

$$F(t) = K_p e(t) + K_i \int_0^t e(\tau)d\tau + K_d \frac{de(t)}{dt}, \qquad (12)$$

where F is the output force command, and K_p, K_i, K_d are non-negative coefficients for the proportional, integral, and derivative terms, respectively. For example, if the desired velocity is greater than the actual velocity, the velocity error becomes positive. Thereby, the proportional component $K_p e(t)$ increases to generate larger force, which increases the actual velocity and reduces the velocity error. The integral and derivative components make the response of the controller more stable and responsive. To further reduce the effects of the coupling dynamics between PBR and pedestal, we set a large mass for the pedestal. A larger mass pedestal with a larger inertia can absorb more energy from the collision, resulting in a smaller collision-caused velocity change.

We developed an automation program to repeat overturning experiments with different single-pulse cosine ground motions. Fig. 5b illustrates the automation workflow. The objective of the automation process is to obtain PBR overturning responses (Eq. 1) to a linear mesh of PGV/PGA and PGA. From the set {(PGV, PGA)}, the automation program popped every pair of (PGV, PGA) and sent it to the controller to actuate a single-pulse cosine ground displacement until the set was empty. To ensure the consistent initial PBR position and orientation across all experiments, the automation program loaded and deleted the PBR model before and after every experiment, respectively. The automation program recorded the PBR states during the overturning experiment, which included the final status of overturned or balanced after the single-pulse ground motions. The automation program was at the highest level in the hierarchical control module. The automation program passed high-level control signals {(PGV, PGA)} to the middle-level motion interpreter and trajectory planner. Then the middle-level trajectory planner passed control signals (velocity commands) to the low-level PID velocity controller, which generated force for the prismatic joint.

With the Bullet physics engine, the VSR supports meshed models of terrains mapped in the real world. For example, Fig. 6 shows the process of creating 3D geometric models of terrain and PBR. The terrain and PBR were mapped by UAS and SfM at a study site of Double Rock, which is located close to both the population center of San Luis Obispo and the critical infrastructure at Diablo Canyon in south-central coastal California (The PBR is named DRE2 in Rood et al., 2020). We first separated the terrain and PBR in the mesh model reconstructed by SfM. The separated terrain, however, lacked supporting and lateral surfaces, which were not reconstructed because of their invisibility without physically removing the PBR. To address this, we completed the missing surfaces by manually adding planes such that the slopes of the added planes were close to the slopes of the local jointing surfaces on the terrain. Similarly, we added planes to the PBR geometric model and built a closed-surface mesh model using Poisson reconstruction, as shown in Fig. 6. From a closed-surface mesh model, we used CAD software (e.g., Autodesk Fusion 360) to measure the mass and moment of inertia of the PBR, which is made of chert with a density of 2,110 kg/m^3. The height of the PBR is approximately 12 m

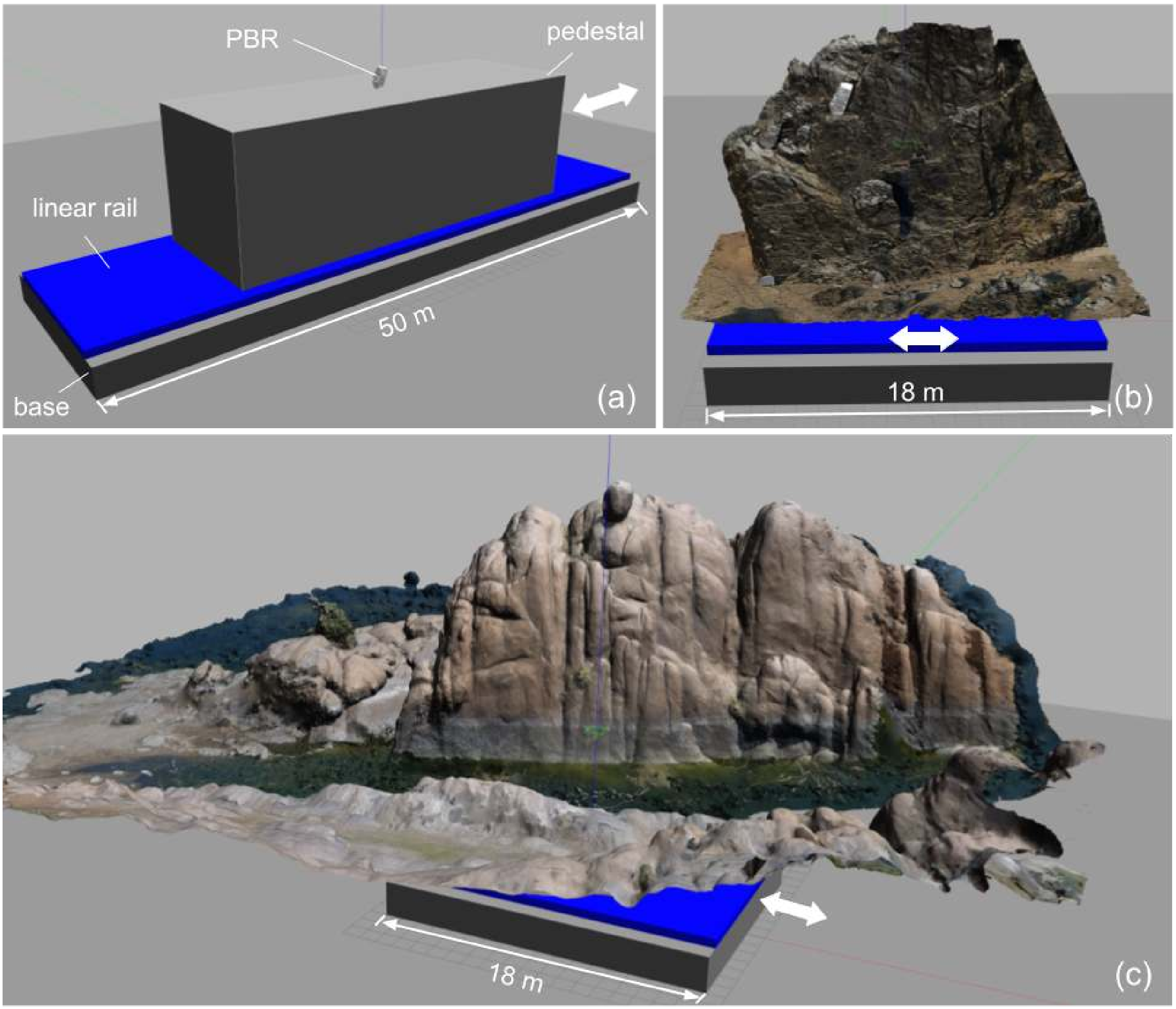


Figure 4 VSR with (a) flat and (b, c) realistic terrains. Unpiloted Aircraft System (UAS) and Structure from Motion (SfM) produced the full-scale realistic terrains in (b) Double Rock, California and (c) Granite Dells near Prescott, Arizona. Arrows indicate ground motion directions.

above the surrounding ground surface.

Geometric simplifications like this, particularly at the base or interface, introduce potentially significant uncertainties in overturning fragility. Complex basal conditions, as is the case for many PBRs, effectively introduce multiple points of rocking or potential uplift. The resulting increase in fragility was evident in the shake table tests of Purvance et al. (2008) and the analytical model of Wittich and Hutchinson (2017). More recent shake table testing by Saifullah and Wittich (2021) quantified that the overturning demand can vary up to $\pm 50\%$ because of small modifications in the basal geometry from surveying techniques. The basal contact can be made arbitrarily complex if necessary in the VSR, but that was not the goal of this paper. Although the geometric modeling approach taken herein does not fully capture the basal interface, this paper aims to provide a demonstration of a first-generation technology for modeling the dynamics of PBRs.

4 Validation

4.1 Velocity Controller

We evaluated the performance of the PID velocity controller in the VSR using a single-pulse cosine displacement ground motion and realistic ground motion. The objective of the PID velocity controller was to generate the translation force to actuate the pedestal (flat surface or realistic terrain), such that the actual velocity measured from the pedestal (ground motion velocity) in the simulation closely matched the desired velocity. Fig. 7a presents an example where the desired velocity and actual velocity from a single-pulse cosine displacement ground motion with PGA of 0.2 g and PGV/PGA of 0.8 s. The overlap of the two velocities in Fig. 7a demonstrates the good performance of the PID velocity controller. We observed similar matches between the desired velocity and actual velocity in other single-pulse cosine displacement ground motions.

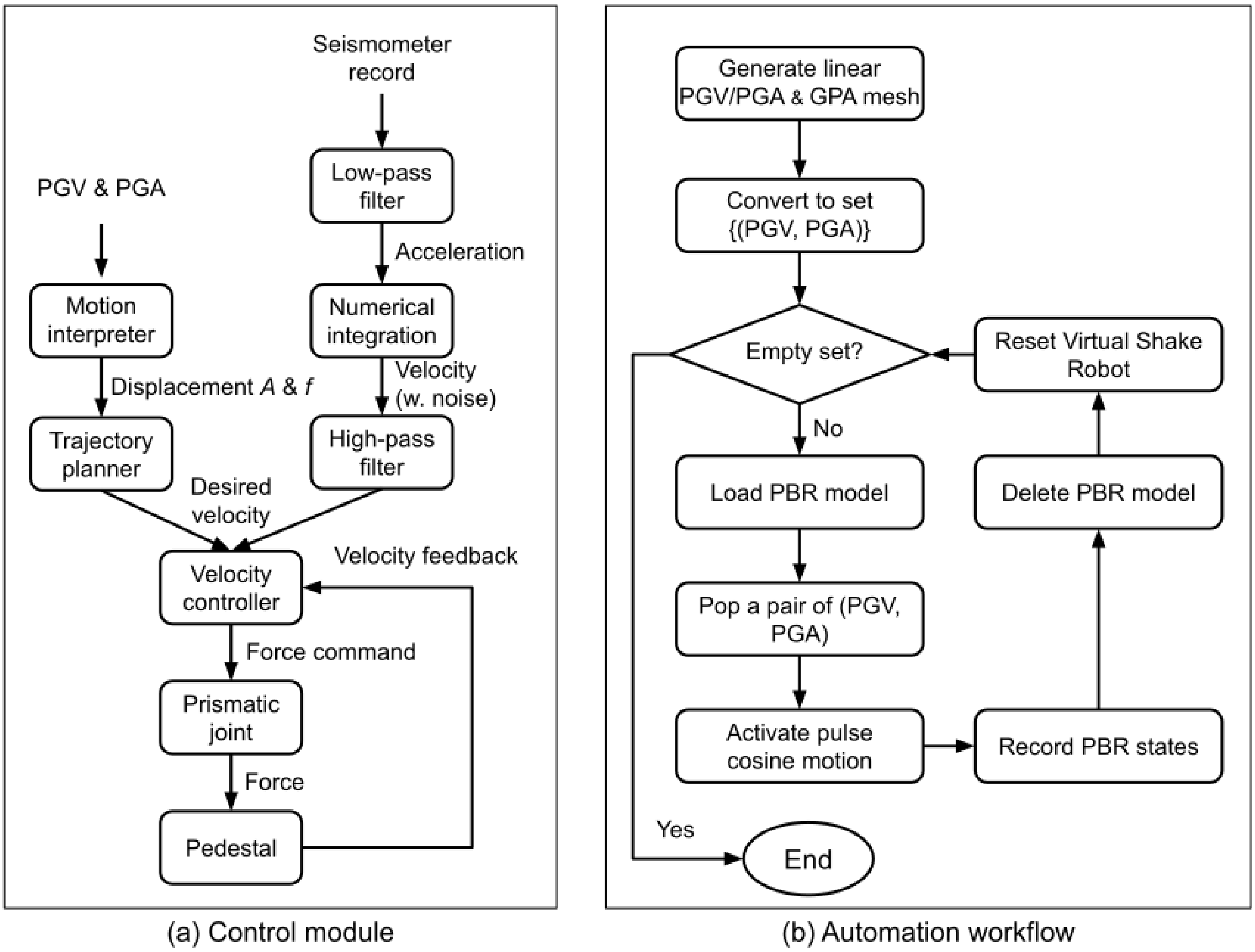


Figure 5 (a) Control module and (b) automation workflow of the VSR.

To evaluate the realistic ground motion, we calculated the desired velocity based on raw acceleration data collected in an accelerometer on a physical shake table. Fig. 7b shows the desired and measured velocities from the VSR. During the shake test, the pedestal experienced a rapid disturbance in the actual velocity, as highlighted in the box in Fig. 7b, caused by the overturned PBR. The PID controller was able to quickly correct the velocity disturbance, demonstrating the robustness of the controller. Additionally, we used this realistic ground motion to shake the Double Rock PBR on a flat surface and on the realistic terrain (in the yaw 0° direction). This realistic ground motion overturned the Double Rock PBR on the flat surface but did not on the realistic terrain with the surrounding pedestals, demonstrating the effects of surrounding pedestals on PBR fragility.

4.2 Previous Overturning Experiments

We validated the overturning dynamics by comparing the shake experiments from the VSR and Purvance et al. (2008). Referencing the known height of a nearby steel I-beam section, we estimated the dimension of a wooden block (labeled W2) from Figure 4 in Purvance et al. (2008) as $5.5 \times 1.1 \times 1.1\ cm$. We modeled a cuboid with the same dimension and a density of 1,500 kg/m^3 (wooden density) in Gazebo. The cuboid was placed on flat terrain with coefficients of Coulomb friction, rolling friction, and restitution of 0.6, 0.6, and 0.2, respectively. The two friction coefficients were selected based on dry rock friction (Byerlee, 1978), and the restitution coefficient of a wood block was measured from a normal drop test (Haron and Ismail, 2012). Using the automation program (Fig. 5b), the VSR conducted 2,500 overturning experiments on a linear mesh space of single-pulse cosine displacement ground motions where PGV/PGA $\in$ [0.05, 0.35] and PGA $\in$ [0.05, 0.5]. For each PGV/PGA, the VSR increased the PGA from 0.05 g to search for the first PGA that overturned the cuboid.

Fig. 8 shows the results from the VSR and Purvance et al. (2008). The blue curve delineates the first overturning PGAs of the cuboid from the VSR experiments. The squares represent the results of the wooden block W2 from the physical overturning experiments by Purvance et al. (2008). The gray-filled region, showing prediction intervals about the fragility contours within which 95% of the overturning responses occur, was calculated from empirical equations based on the square data (Purvance et al., 2008). The results from the VSR were within the 95% prediction intervals when PGV/PGA is greater than 0.08 s. When PGV/PGA is

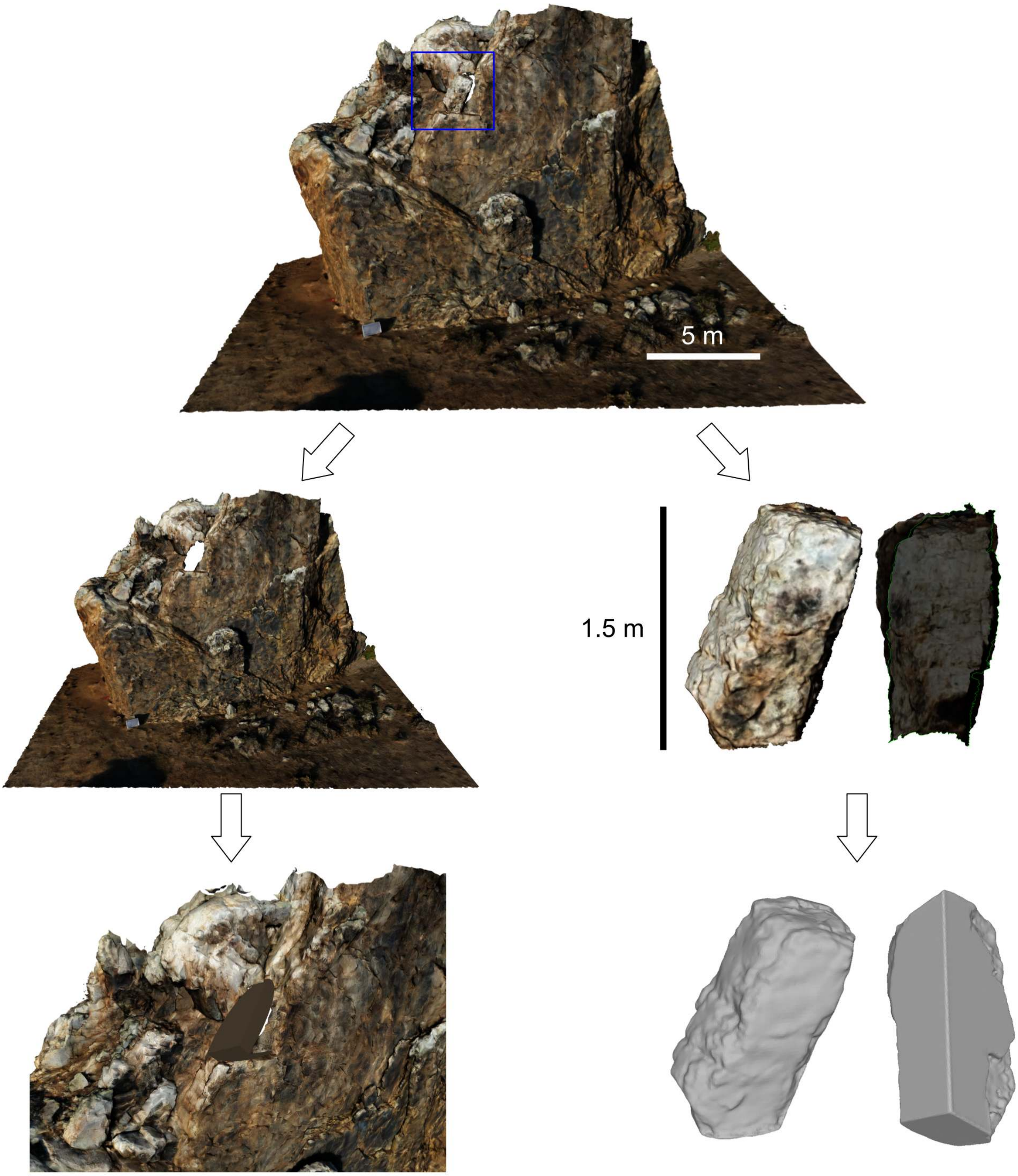
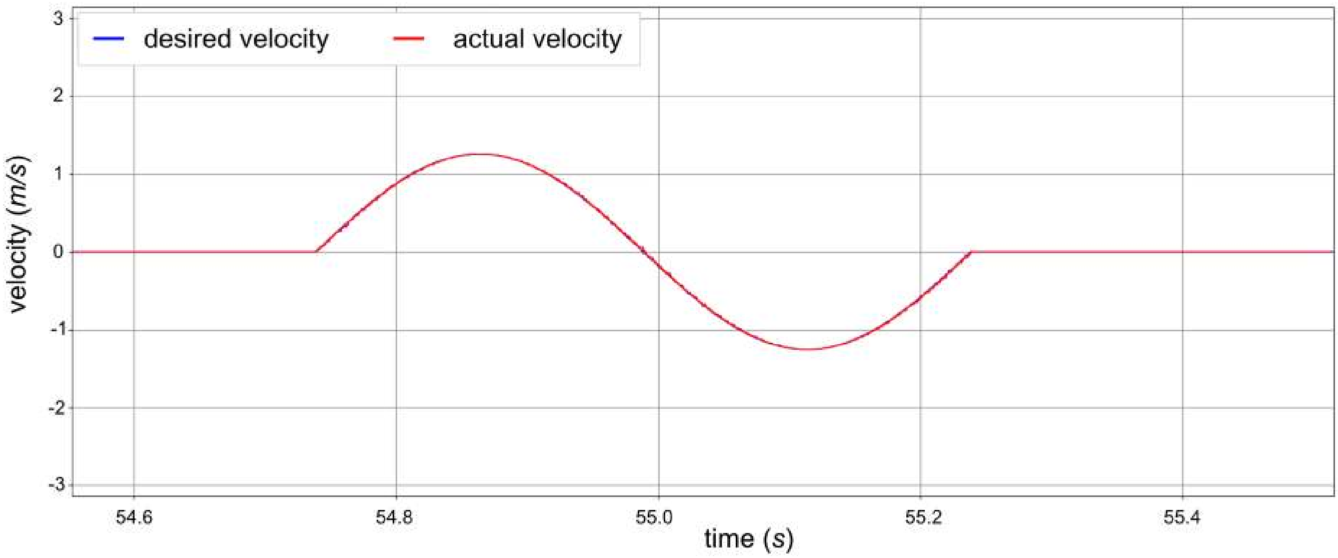


Figure 6 3D geometric modeling of terrain and PBR mapped by UAS and SfM. The site of Double Rock is located in south-central coastal California.

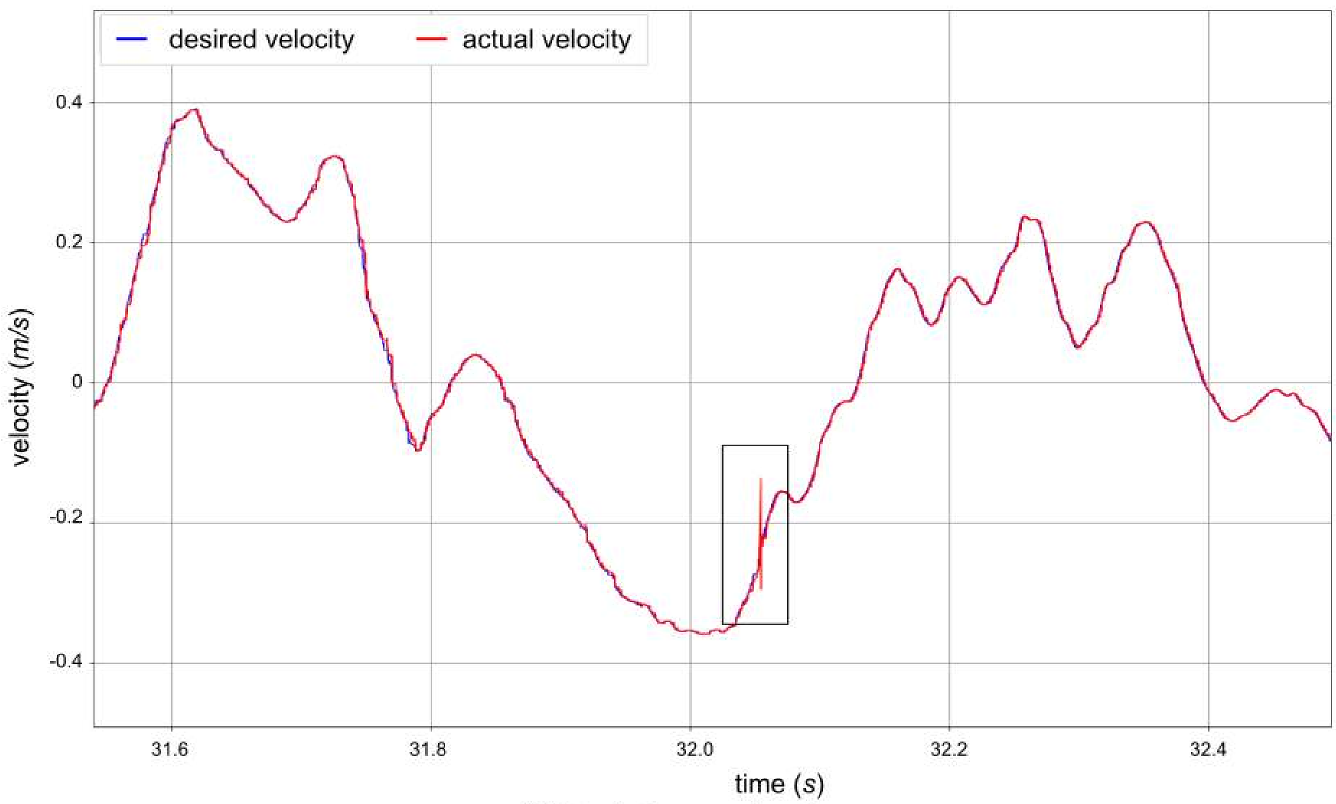
smaller than 0.08 s, the 95% prediction intervals were below the VSR curve (Fig. 8), indicating that the cuboid in the VSR was less likely overturned by ground motions with small periods. Note that the VSR curve was resulted from 2,500 overturning experiments with different parameters densely sampled from the ground motion space. However, the 95% prediction intervals were calculated based on a small number of data points (the squares), and only one data point was collected for PGV/PGA smaller than 0.08 s. Future work should collect more data points to examine the 95% prediction intervals, especially on the small PGV/PGA space.

4.3 Mini shake robot

We developed a physical, mini shake robot to validate the simulation (Chen et al., 2022), as shown in Fig. 9 The mechanical hardware design of the mini shake robot



(a) Single-pulse cosine displacement ground motion



(b) Realistic ground motion

Figure 7 Velocity plots from (a) single-pulse cosine displacement ground motion and (b) realistic ground motion. The PID velocity controller generates translational force to actuate the pedestal of the VSR based on the desired velocity from a trajectory planner or a record of a real-world accelerometer. The actual velocity is measured from the pedestal of the VSR. The desired velocity and actual velocity are overlain in panel (a). (b) A sudden disturbance in the actual velocity, as highlighted in the box, was caused by the overturned PBR but quickly rectified by the PID controller.

adopts a closed-loop stepper motor for actuation and a toothed belt for transmission. We developed the software of the mini shake robot based on ROS, which had a similar robotic software architecture as the VSR (Fig. 3). The mini shake robot uses the same high-level and middle-level control programs as the VSR. The primary difference is the low-level velocity controller: the mini shake robot employs a closed-loop stepper motor with toothed belt transmission, whereas the VSR uses a prismatic motor (prismatic joint). The closed-loop stepper motor consists of a regular stepper motor (specifically, NEMA 34HS31) and an encoder that measures the shaft orientation, forming the feedback loop for the low-level PID controller. With a toothed pulley of 53.98 mm outside diameter, the stepper motor enables the bed to reach a maximum horizontal acceleration of 1.2 g and a

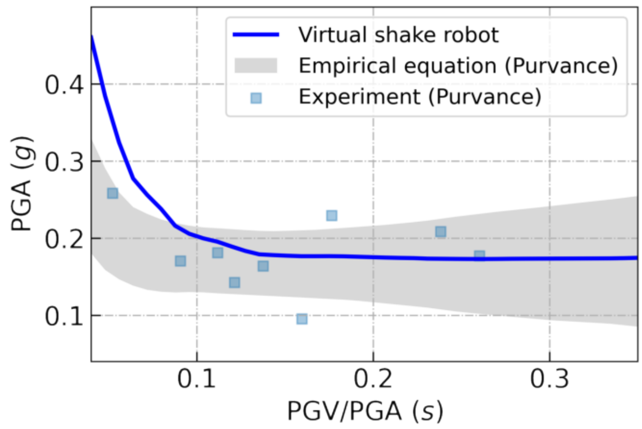
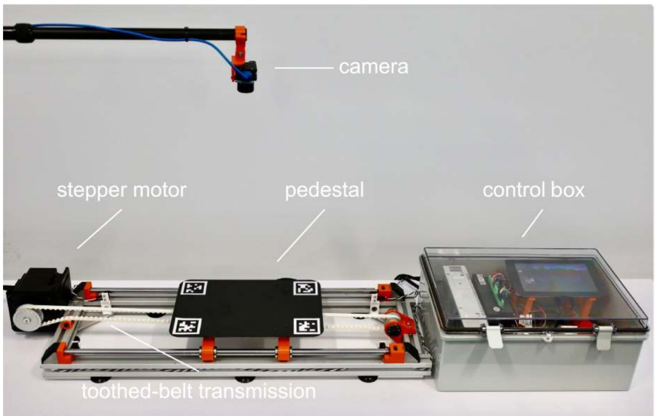
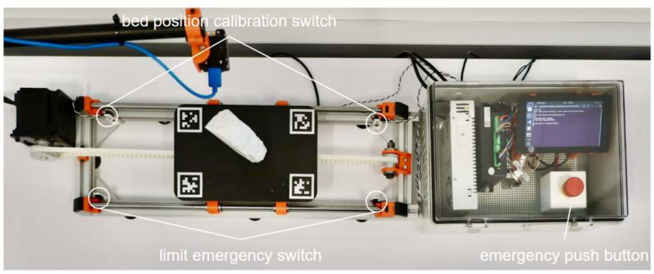


Figure 8 Overturning experiment results of a wooden block (W2 from Purvance et al., 2008). Squares and curve indicate the experimentally observed PGA at which the block overturns for the first time, as PGA is gradually increased from a small value for each PGV/PGA. Gray-filled region delineates prediction intervals about the fragility contours within which 95% of the overturning responses occur.

maximum velocity of 0.5 m/s with up to 2 kg PBR payload. Overall, the mini shake robot provides a low-cost, open-hardware, and open-software platform for earthquake research and education.



(a) Mini shake robot side view



(b) Mini shake robot top-down view

Figure 9 Mini shake robot from (a) side view and (b) top-down view. (b) 3D-printed PLA PBR is placed on the pedestal. Reprinted from Chen (2022) with permission.

The mini shake robot provides a reverse method for simulation validation. Using the mini shake robot, we conducted experiments with small-scale, free-standing blocks such as 3D-printed PBRs. We then repeated these experiments in simulation using the VSR to compare the physical and simulation results. Specifically, we downscaled the Double Rock PBR from the height of 151.0 cm to 12.8 cm and used polyethylene terephthalate glycol (PETG) material with a density of 1,240 kg/m^3 to 3D print the PBR. The PETG PBR weighted 403.5 g. Following the instructions in Anooshehpoor et al. (2004), we applied grip tapes on the bed to increase friction (no sliding friction during the experiments). Previous studies used a crane machine to lift and rest a heavy PBR after each overturning experiment (e.g., Saifullah and Wittich, 2021). Because of the small size and light weight of the PETG PBR, we were able to reset its pose precisely without the use of a crane. Based on the geometry and physical properties of the PETG PBR, we created a free-standing block with the same dimension in simulation and validated its overturning dynamics using the VSR. Fig. 10 presents the overturning validation results. We used the logistic regression model (Purvance, 2005) to approximate the boundary curves between the overturned and balanced responses. Despite the observation that the real PETG PBR was more fragile, the boundary curves from the simulation and physical experiments were close.

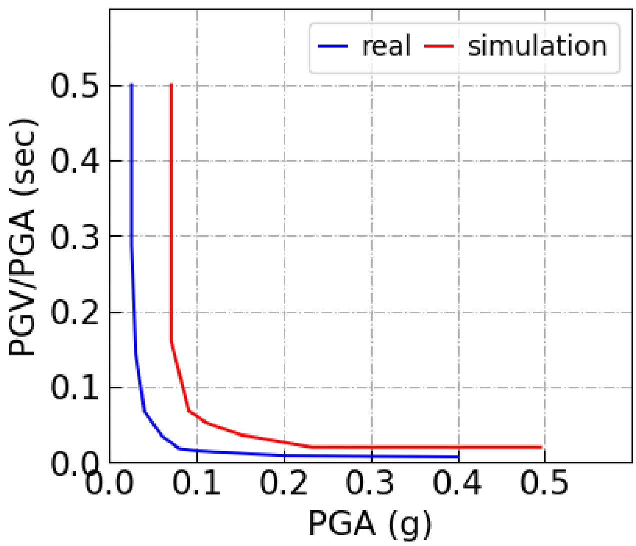


Figure 10 Overturning response diagram of 3D-printed Double Rock PBR. Blue curve represents logistic regression results from the mini shake robot. Red curve represents logistic regression results from the VSR. Reprinted from Chen (2022) with permission.

In the second experiment, we downscaled the Double Rock PBR to a height of 12.0 cm and 3D-printed it with polylactic acid (PLA) material, which has a density of 1,250 kg/m^3, resulting of a weight of 105 g. To validate fragility anisotropy, we placed the PLA PBR with two initial orientations (yaw angles of 0° and 270°) on the bed. Using the mini shake robot, we obtained the response diagrams from a set of ground motions, as shown in Fig. 11. The response diagrams show that PLA PBR oriented at a yaw angle of 0° is more fragile than at yaw angle of 270°. This result is consistent with analysis from the previous studies, which found PBRs with smaller minimal contact angle along the motion direction is more fragile (Purvance et al., 2008; Haddad et al., 2012). The resulting boundary curves have a similar pattern to the simulation results of the Double Rock PBR with original dimensions and chert density (see Section 5.2).

5 Experiments

We conducted a set of experiments to demonstrate the applications of the VSR, test the limitations of certain PBR assumptions, and investigate the PBR overturning and large-displacement processes.

5.1 Cuboids Overturning on Flat Terrain

Because rectangles were typically studied in the previous overturning studies (Milne and Omori, 1893; Kirkpatrick, 1927; Housner, 1963; Yim et al., 1980; Purvance et al., 2008), we examined the overturning dynamics of cuboids on a flat pedestal using the VSR. Specifically, we created four cuboids with dimensions of $1{\times}1{\times}2\ m$, $0.5{\times}0.5{\times}1\ m$, $1{\times}1{\times}3\ m$, and $1{\times}2{\times}3\ m$ (Fig. 12a). These four cuboids had the same densities of 2,110 kg/m^3 (chert density), 0.6 coefficient of Coulomb friction (dry rock friction), 0.6 coefficient of rolling friction (dry rock friction), and 0.38 coefficient of restitution (based on rockfall energy loss reported in Dorren, 2003).

Fig. 12b-e shows the overturning response to single-pulse cosine ground motions. In the cuboid's coordinates, the horizontal ground motions were along the x axis (as the red arrow indicated in Fig. 12a). By comparing Fig. 12b and c, the cuboid with a larger height-width ratio was more fragile. With the same height-width ratio, the smaller cuboid was more fragile by comparing Fig. 12b and d. These findings are consistent with the previous reports on rectangle experiments on a 2D plane (Purvance et al., 2008; Anderson et al., 2014; Yim et al., 1980). The cuboids with 1×1×3 meters and 1×2×3 meters had similar overturning response diagrams except for a small number of noises (Fig. 12c, e), which suggested the overturning was not affected by the cuboid materials on the y axis (as the green arrow indicated in Fig. 12a). For an object with more complex geometry, however, extending materials on the y-axis is anticipated to lead to a more intricate overturning response, such as twisting behaviors, which would need further investigation to be confirmed.

5.2 PBR Fragility Anisotropy on Flat Terrain

Because geometries of natural PBRs are often asymmetrical, the overturning responses should vary from different ground motion directions–fragility anisotropy. However, most previous studies simplified PBR geometries and only considered the minimal contact angle with the pedestal (e.g., Purvance et al., 2008; Haddad et al., 2012). Veeraraghavan et al. (2016) studied fragility anisotropy based on a rigid-body dynamics algorithm (Chapter 2 in Veeraraghavan, 2015). Using the VSR, we examined the fragility anisotropy of the Double Rock PBR (Fig. 6) simply by placing the PBR on a flat pedestal in 12 different initial orientations (spaced every 30° on yaw). The orientation of PBR was defined in Fig. 13a. By placing the PBR with different orientations, we simulated the varying ground motion directions. We set the Double Rock PBR with the same physical properties as the cuboids described above and applied the automation program (Fig. 5b).

Fig. 13b-f depict the overturning response diagrams for the fragility anisotropy study on a flat pedestal. From the opposite directions such as Fig. 13b, d or Fig. 13c, e, the same single-pulse cosine ground displacements produced different overturning responses. By comparing Fig. 13b-d, the PBR was less fragile to the 90° ground motions than 0° or 180° ground motions. In Fig. 13f, some balanced responses separate a small cluster of toppled responses on the left (green polygon) and the major boundary curve. For this ground motion direction (yaw 300°), the logistic regression method proposed by Purvance (2005) would not be the ideal model to compute the overturning probability, because the logistic regression model, calculated based on the data near the first PGA for a PGV/PGA, would fail to take into account the balanced responses between the two clusters of toppled responses.

5.3 PBR Fragility Anisotropy on Realistic Terrain

Using the VSR, we investigated the effects of surrounding pedestals on the overturning responses of the Double Rock PBR. Most previous studies on the dynamics of free-standing blocks assumed that the blocks had no interaction with adjacent objects. Konstantinidis (2008) explored the overturning dynamics of a free-standing block anchored to a wall via chains. Bao and Konstantinidis (2020) investigated 2D analytical models to study the dynamics of a free-standing block considering impact with an adjacent wall. However, no previous PBR studies have directly accounted for lateral supports in 3D. In this experiment, we set up the terrain model in different orientations to study fragility anisotropy with surrounding pedestals (Fig. 14a, c). The Double Rock PBR had the same physical properties as described earlier. We set the same contact parameters (friction and restitution) for the terrain model.

Fig. 14b and d show the overturning response diagrams of the Double Rock PBR on the realistic terrain model. Note that the ground motion ranges in Fig. 14 are larger than those in Fig. 13. By comparing Fig. 13b and Fig. 14b, the surrounding pedestals significantly reduced the PBR fragility, when the ground motion was along the yaw 0° direction. From Fig. 13c and Fig. 14d, the surrounding pedestals only slightly reduced the PBR fragility, when the ground motion was along the 90° direction. The effects of the surrounding pedestals varied for different ground motion directions.

5.4 PBR Large Displacement on Realistic Terrain

In this experiment, we investigated the large-displacement dynamics induced by ground motions. Using the VSR, we recorded the PBR states during the experiments of PBR fragility anisotropy on the Double

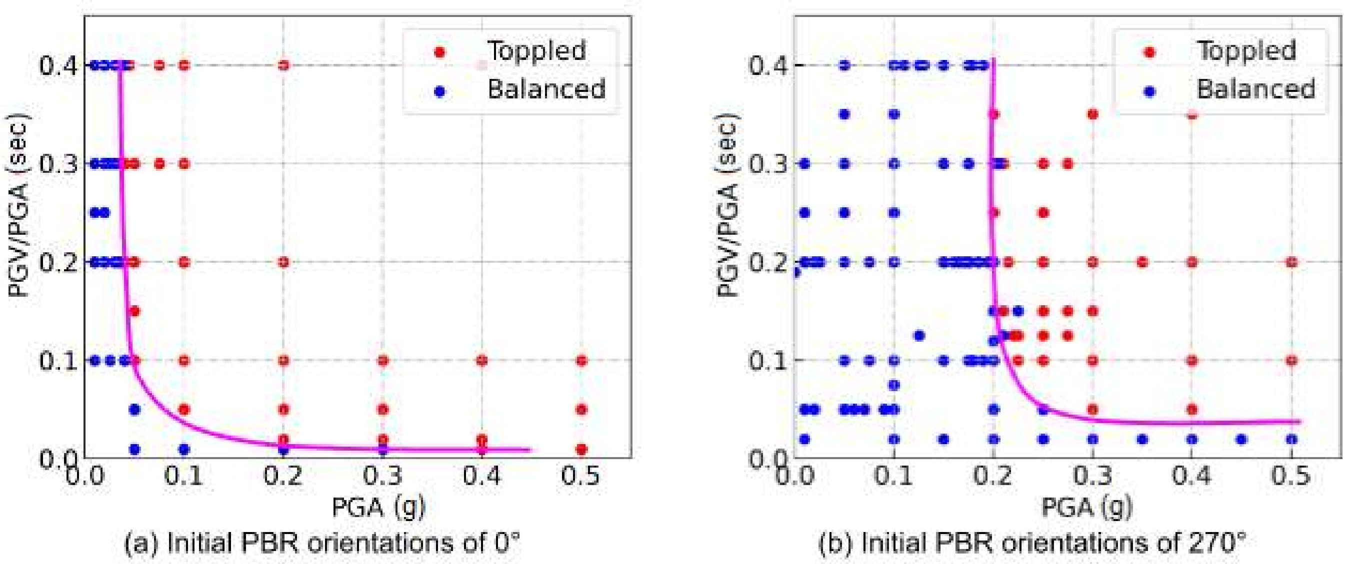


Figure 11 Overturning response diagrams of 3D printed Double Rock PBR from initial orientations of yaw 0° and 270°. The red and blue dots represent overturning responses of being toppled and balanced after single-pulse cosine ground motions, respectively. The curves indicate the boundaries of the overturning responses. Reprinted from Chen (2022) with permission.

Rock terrain (Section 5.2). Fig. 15 presents the results of the PBR large-displacement experiments for the ground motion in the yaw 0° direction. Fig. 15a shows a time-lapse snapshot of a PBR trajectory. Fig. 15b illustrates all 692 trajectories of the PBR after toppled, compared with Fig. 14b that includes the overturned or balanced responses to ground motions. Fig. 15c-f depict the trajectories of the PBR in different PGA ranges.

Using the recorded trajectories, we analyzed the relationship between ground motions and large displacements. As shown in Fig. 15c-f, the number of trajectories increases with PGA (increase in overturned PBRs). Comparing the trajectory plots (Fig. 15c-f) reveals how trajectory and velocity change with PGAs. Fig. 16 highlights the relationship between trajectory lengths with ground motions. Concurrently large PGA and PGV/PGA result in a long trajectory. Only one large value in either PGA or PGV/PGA is insufficient to produce a long trajectory. Fig. 16b, c show an increasing trend of trajectory length as PGA or PGV/PGA increases. The trajectory data points around 6 m (Fig. 16b, c) correspond to situations where the overturned PBR lands on the niche position shown as the white box in Fig. 15a.

We conducted a correlation analysis to quantify the relationship between ground motions and large displacements. The correlation analysis used the 692 trajectories resulting from the overturned states (red data points in Fig. 14b). Each trajectory is characterized by its trajectory length, largest velocity, and terminal distance. The terminal distance is the Euclidean distance between the start and terminal positions. For each PGA, we ensembled the trajectories from all corresponding PGV/PGAs (all red data points on a PGA column in Fig. 14b). Similarly, for each PGV/PGA, we ensembled the trajectories from all corresponding PGAs. In each ensemble, we calculated the trajectory statistics such as mean trajectory length, mean largest velocity, and mean terminal distance. Fig. 17 plots the trajectory statistics and ground motions. The PGA and PGV/PGV positively correlates with mean trajectory length, mean largest velocity, and mean terminal distance, as the R^2 and p value summarized in Table 1. The correlation analysis results reject a null hypothesis that the trajectory statistics and ground motions are uncorrelated.

6 Discussion

6.1 Validation

The validation experiments examined the performance of the VSR in terms of the velocity controller and overturning dynamics. The disturbance in the actual velocity in Fig. 7b shows the coupling dynamics between the PBR and pedestal. However, the effects of such coupling dynamics have often been neglected in previous studies. The quick correction in the actual velocity (Fig. 7b) indicates that the PID controller was robust to such disturbance. The overlay between the desired velocity and actual velocity in Fig. 7 shows the good performance of the PID velocity controller. Additionally, the realistic ground motion (0.4 g PGA and 1.2 s PGV/PGA) overturned the Double Rock PBR on the flat surface but did not on the realistic terrain with the surrounding pedestals, which agrees with the overturning response diagrams in Fig. 13b and Fig. 14b.

We compared the overturning results from the VSR and Purvance et al. (2008). When PGV/PGA was greater than 0.08 s, the overturning results from the VSR were consistent with Purvance et al. (2008). The reasons for the difference from the low PGV/PGAs remain to be explored. However, we observed that the first overturning PGAs in the VSR were less scattered from the median fragility contour than Purvance et al. (2008). More physical experiment data points in the low PGV/PGA region are needed to refine the 95% bounds. Additionally, low PGV/PGA motions typically exhibit higher-frequency

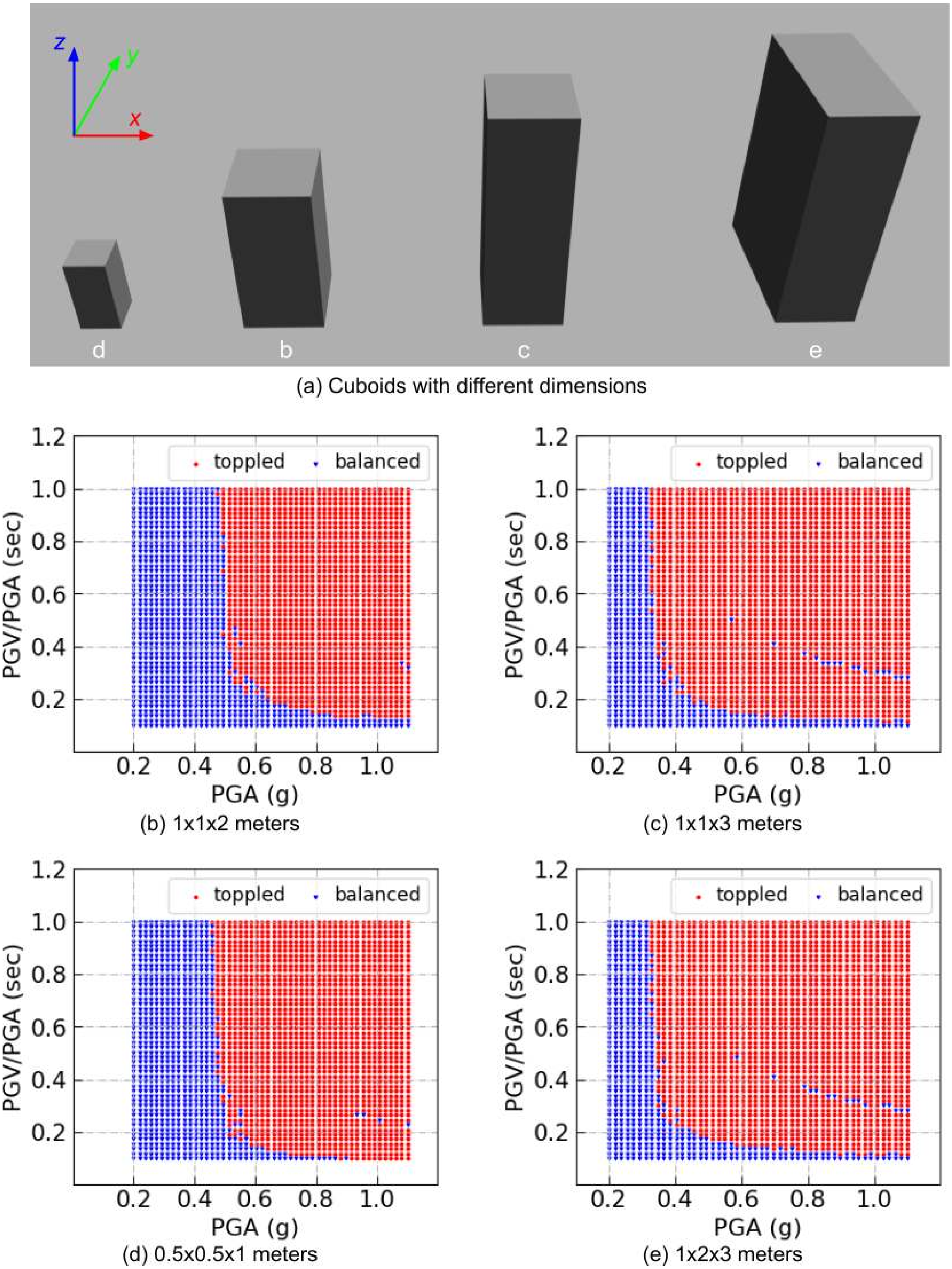


Figure 12 Overturning response diagrams from cuboids with different dimensions on a flat pedestal. Ground motions are along x direction in the $x \times y \times z$ dimension. Red and blue dots represent overturned response and balanced response, respectively. Horizontal axis represents the peak ground acceleration (PGA) in gravity constant. Vertical axis represents the ratio of peak ground velocity to peak ground acceleration (PGV/PGA). Unit of PGV/PGA is second. The blue dots within the toppled zones represent the PBRs being pulled back to balanced states by the returning ground motions.

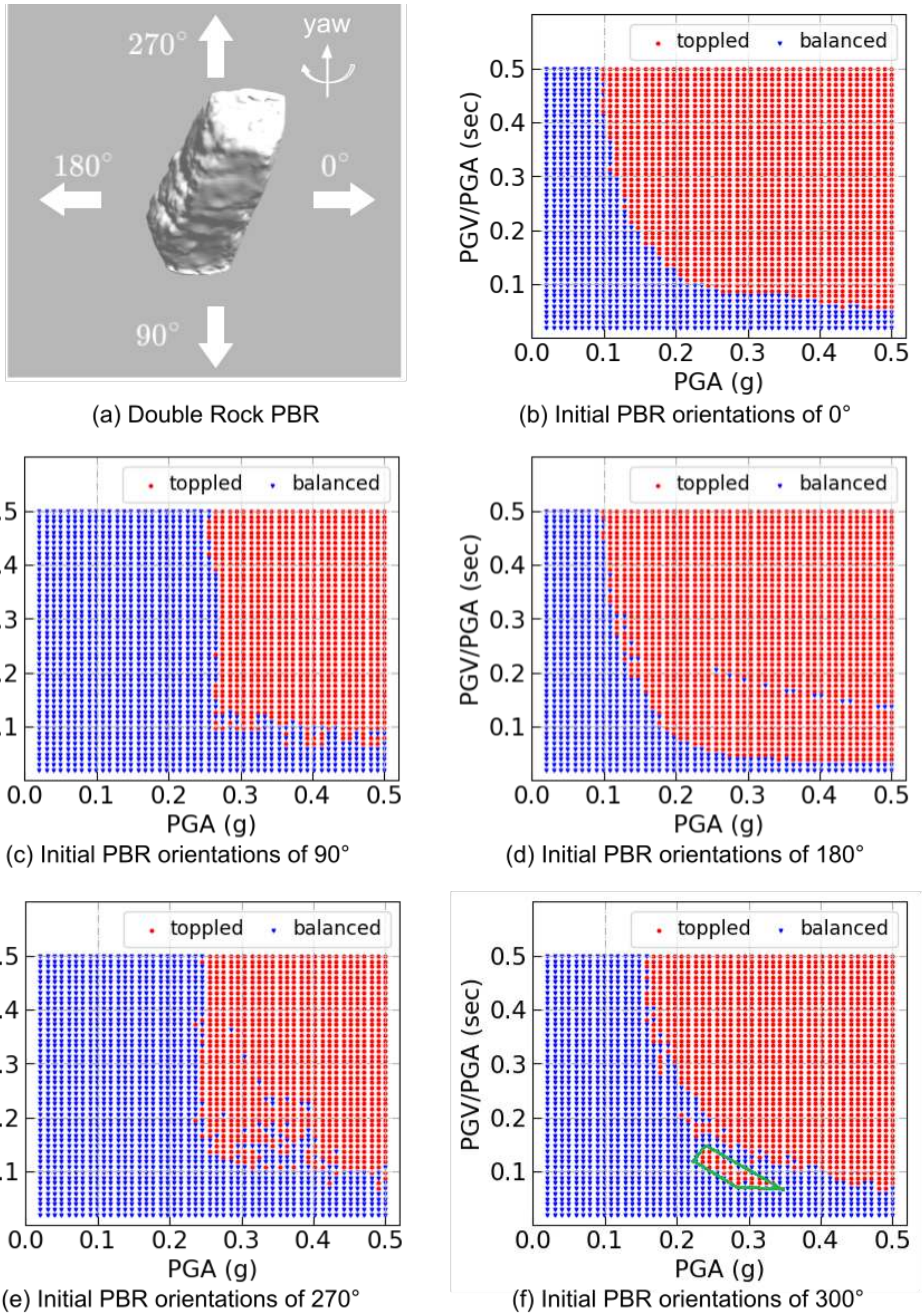
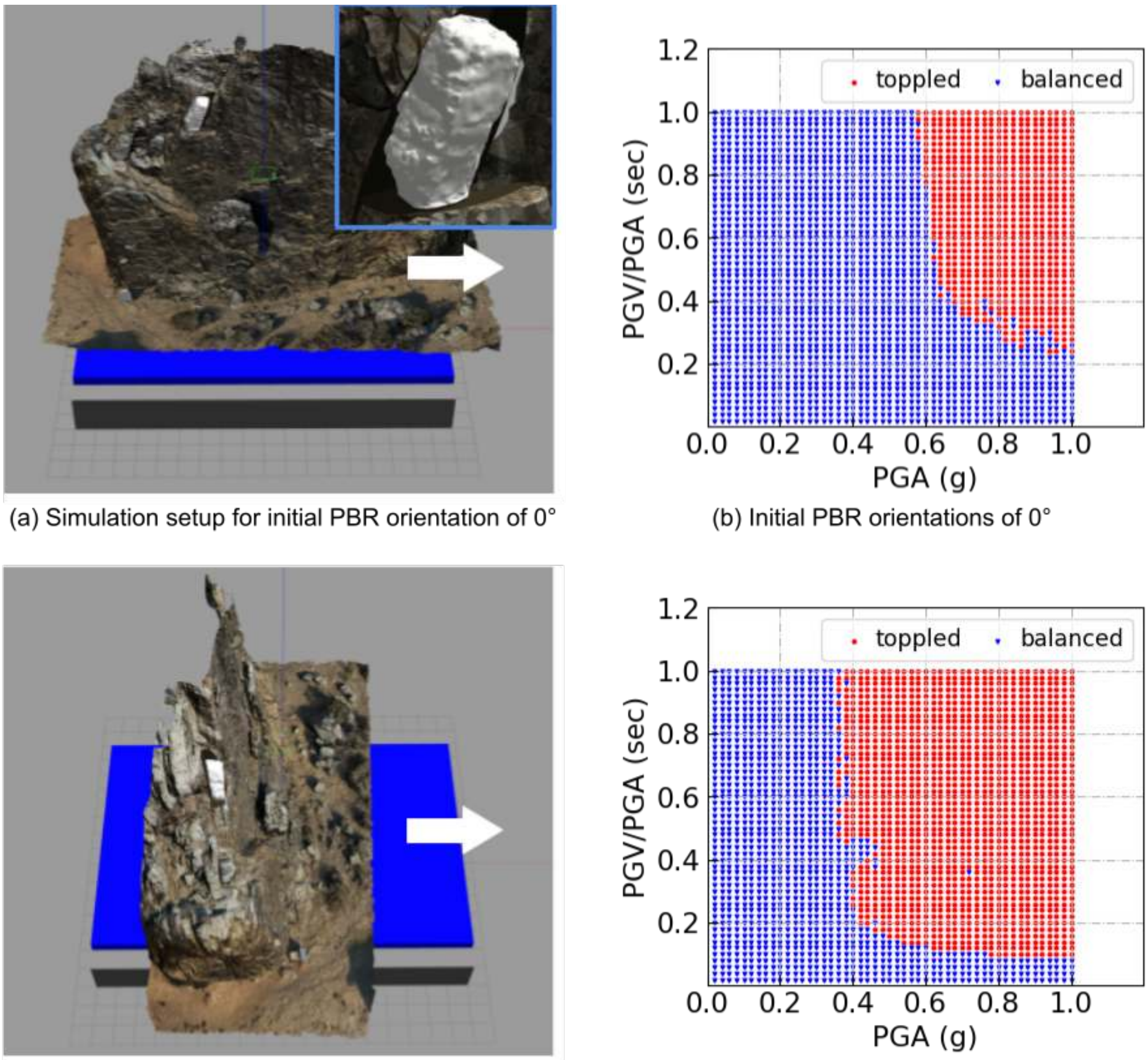


Figure 13 Overturning response diagrams from Double Rock PBR on a flat pedestal. (a) Arrows indicate the equivalent ground motion directions of the different initial PBR orientations.

movements. On physical free-standing structures, high-frequency movements may trigger a slight rocking or wobbling because of imperfections in the basal contact. Because rocking behavior is nonlinear to orientation, even such a slight rocking from higher-frequency movements can increase susceptibility to overturning, especially when compared to VSR predictions that do not account for physical imperfections at the base. Thus, validating the rocking behaviors resulting from high-frequency movements would require more physical ex-



(a) Simulation setup for initial PBR orientation of 0°

(b) Initial PBR orientations of 0°

(c) Simulation setup for initial PBR orientation of 90°

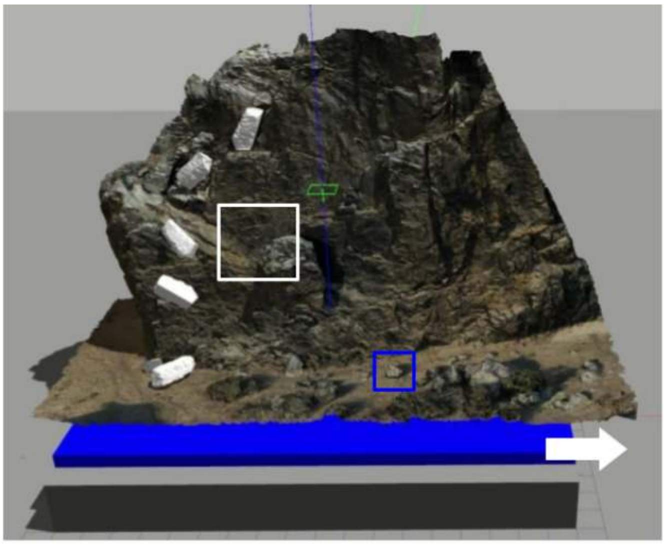
(d) Initial PBR orientations of 90°

Figure 14 Overturning response diagrams from Double Rock PBR on the realistic pedestal mapped from UAS and SfM. (a, c) Arrows indicate directions of the single-pulse cosine ground displacement motions. Inset at (a) shows a zoom-in of the PBR on the realistic terrain.

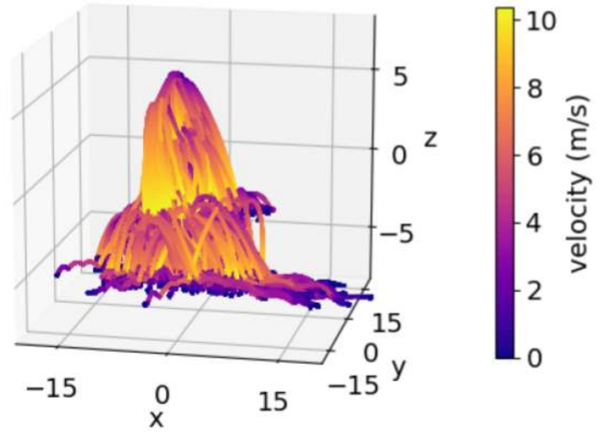
periments.

We built the mini shake robot to examine the performance of the VSR on PBR overturning dynamics. Because of using ROS, the mini shake robot reused the control software from the VSR, ensuring consistent ground motion generation processes in the simulation and physical experiments. In the first validation experiment, the boundary curves from the simulation and physical experiments were close. We noticed that the real PETG PBR was more fragile and are working to identify the causes of the difference. During the experiment, we observed high-frequency mechanical vibrations along the vertical direction (perpendicular to the ground motion direction) on the mini shake robot. Additionally, because the PETG PBR was downscaled and we know that smaller PBRs are more fragile (see Section 5.1), the fragility of the downscaled PBR may be more easily affected by ground motion noise. Therefore, this scale concern warrants further experimentation to validate the VSR using full-scale testing and rock material. To verify the consistency of the overturning results regardless of the PBR scale, future work should repeat the same experiment with 3D printed models at various scales. Because iterative algorithms for solving MLCP may not always guarantee convergence or unique solutions, resulting in variations in rock responses (Veeraraghavan et al., 2020), future work should investigate solutions to mitigate such uncertainties. For example, following Purvance et al. (2008), we can use the Monte Carlo method to construct a probabilistic model and to quantify the uncertainties.

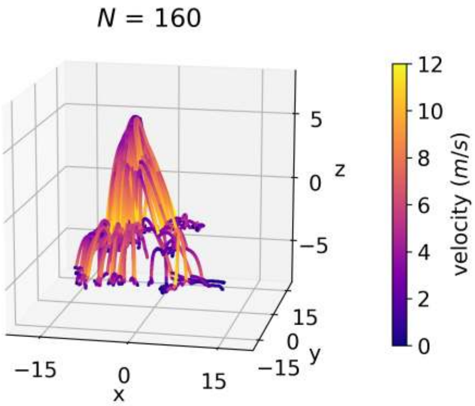
In the second mini shake robot experiment, the fragility anisotropy results from the mini shake table were consistent with the results from the VSR, providing strong evidence of qualitative consistency in fragility analysis. The same pattern of boundary curves was observed in both simulation and physical experiments—the initial PBR orientation of 0° was more fragile than 270°. Given that the edge of the Double Rock PBR appeared well-defined and perpendicular to the



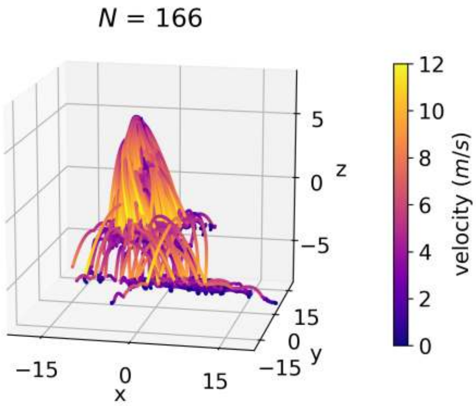
(a) Photo of time-lapse snapshots from a single ground motion



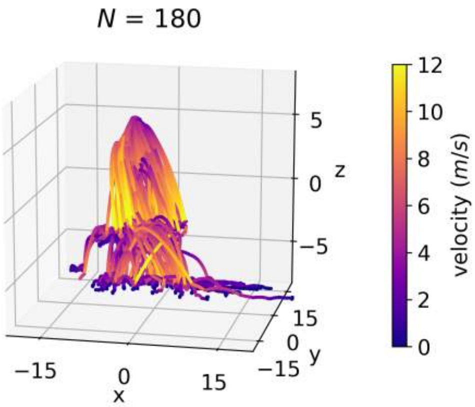
(b) Scatter plot of PBR trajectories from all ground motions



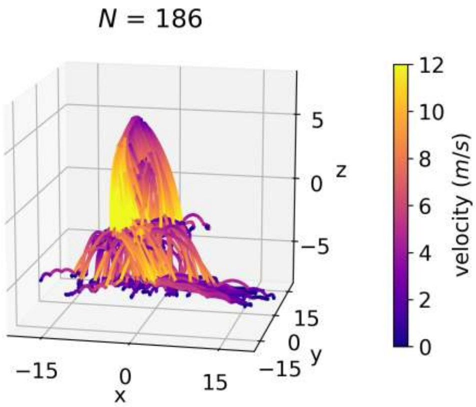
(c) PGA ∈ (0.5, 0.7]



(d) PGA ∈ (0.7, 0.8]



(e) PGA ∈ (0.5, 0.7]



(f) PGA ∈ (0.7, 0.8]

Figure 15 Large-displacement experiment on the Double Rock site from ground motions in the yaw 0° direction. Arrow in panel (a) shows the direction of the single-pulse cosine ground displacement motions. The PBR trajectories in panel (b) are relative positions to the terrain. Colors indicate absolute velocity of the PBR. (c-f) PBR trajectories from different PGA ranges, where N represents the number of trajectories.

ground motion direction (yaw 270°), one might expect only rocking behaviors. However, we observed complex motions such as twisting, point uplift (rocking on a corner), planar uplift (rocking on an edge), and rock-twisting (twisting while in an uplifted state), in both the simulation and experimentation. This complexity arises because rock behaviors are affected not only by the contact conditions but also by factors such as the

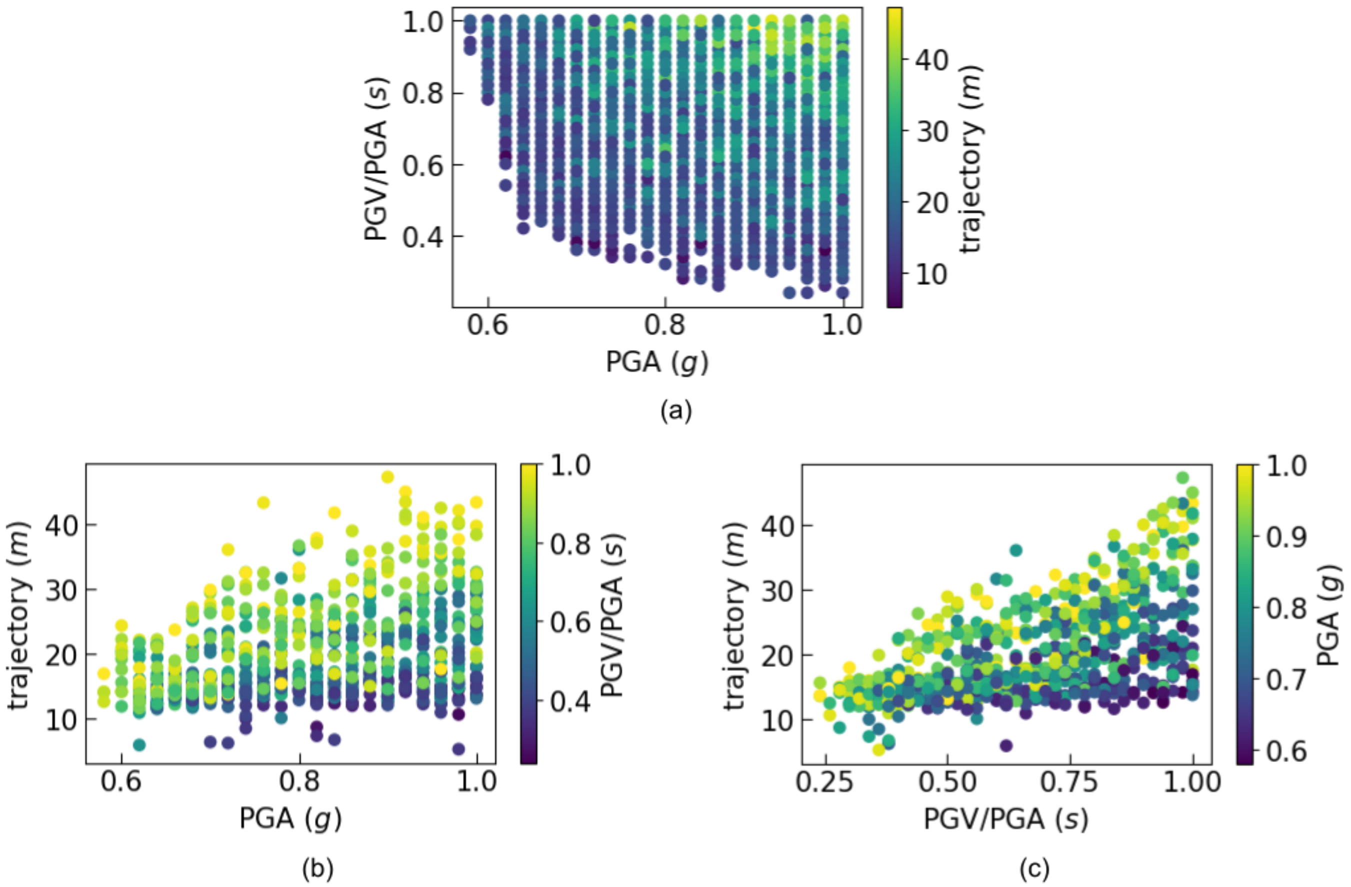


Figure 16 PBR trajectory lengths and ground motions. Data points represent 692 trajectory lengths of the Double Rock PBR after being toppled from ground motions. (a) Trajectory lengths for PGA and PGV/PGA are represented by color-coded dots. (b) Color dots indicate PGV/PGA for PGA and trajectory length. (c) PGA for PGV/PGA and trajectory are displayed by color dots.

R^2	PGA	PGV/PGA
mean trajectory	0.90	0.92
mean largest velocity	0.82	0.78
mean terminal distance	0.73	0.87

p-value	PGA	PGV/PGA
mean trajectory	2.3×10^{-11}	3.2×10^{-22}
mean largest velocity	5.5×10^{-9}	1.0×10^{-13}
mean terminal distance	3.5×10^{-7}	3.9×10^{-18}

Table 1 Large displacement statistics and ground motion correlation analysis

mass center, inertia matrix, and geometry. In our experiments, it took only a few seconds up to a minute to finish one overturning experiment in the VSR. In DEM, the computational time increases with the complexity of geometry, discretization, and contact stiffness. Finishing one overturning experiment in DEM usually takes a few minutes to hours. The rapid deployment of the VSR facilitates qualitative analysis of PBR fragility, which aids in field data collection and assessment such as searching for the most fragile PBRs in the field.

6.2 Overturning and Large Displacement Studies

The VSR offers an integrated solution for studying the dynamics of both overturning and large-displacement processes. From the overturning experiments of cuboids, the relationships between the dimension parameters and fragility were consistent with the previous analytical solutions (Purvance et al., 2008; Anderson et al., 2014). The VSR advances the study of PBR overturning dynamics in many aspects. For example, the overturning experiments indicated that the PBR overturning dynamics were more complex than previously understood. Specifically, the PBR overturning responses were found to vary from ground motion directions and the presence of surrounding pedestals.

Most of our experiments employed single-pulse cosine displacement ground motions. First, such ground motions were easy to synthesize, enabling us to comprehensively cover the configuration spaces of PGA and PGV/PGA. Second, while PGA and PGV/PGA are commonly used as representative ground motion descriptors in PBR studies, recent concerns have emerged regarding their use in PBR studies. When two seismometer data records share the same PGA and PGV/PGA values but differ in other properties, they may produce different PBR overturning responses. To mitigate such ambiguity, we adopted single-pulse cosine displacement ground motions.

The VSR's ability to support realistic terrain models allowed for the seamless study of large-displacement

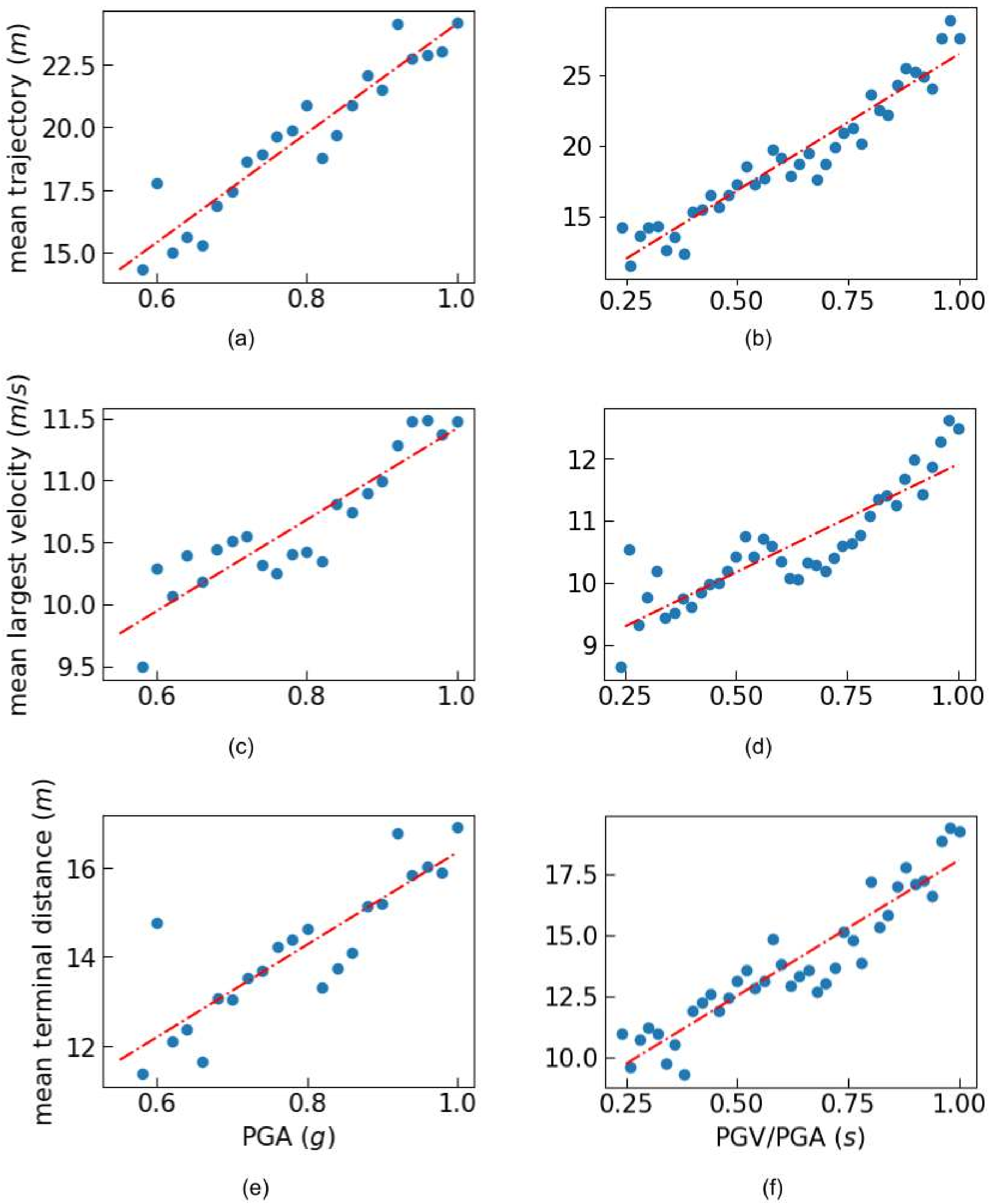


Figure 17 Correlation analysis between large displacement statistics and ground motions. (a, c, e) Correlation between PGA and mean trajectory length, mean largest velocity, and mean terminal distance. (b, d, f) Correlation between PGV/PGA and mean trajectory length, mean largest velocity, and mean terminal distance. Red dash-dotted lines result from least-square linear regression.

dynamics. This functionality of large-displacement analysis enabled the trajectory prediction of overturned PBRs, with potential applications in rockfall hazard zoning and the study of rocky slope development. Additionally, the large-displacement analysis revealed the relationship between large displacement statistics and ground motions for PBRs. The PGA and PGV/PGV positively correlated with mean trajectory length, mean largest velocity, and mean terminal distance. When the overturning dynamics provide upper-bound ground motion constraints by studying fragile PBRs, the large displacement dynamics provide lower-bound ground motion constraints by studying overturned PBRs. For example, given a known trajectory of an overturned PBR, lower bounds of PGA and PGV/PGV can be inferred from Fig. 17b, c. Together, the overturning and large displacement dynamics form complementary methods to refine ground motion estimation.

6.3 Limitations and Future Work

In future research, we will investigate several critical aspects to enhance both the VSR functionality and scientific insights. First, the previous study by Veeraraghavan (2015) demonstrated that 3D PBR fragility results are more sensitive, indicating higher fragility, compared to their 2D counterparts. Looking ahead, the development of the VSR that incorporates 3D ground motions will be a significant advancement beyond the existing 1D ground motion assessment methods. Second, our future work should align with the workflow delineated by Rood et al. (2020), focusing on the exploration of additional PBRs within the Double Rock site. Expanding our scope to encompass a broader array of PBRs has the potential to enhance the accuracy of ground motion estimation. Because of the significant role of contact geometry in rock response, future work should thoroughly investigate the VSR's ability to model dynamic processes involving complex contact surfaces. Such modeling would include complex interface geometry, variable properties and rheology, and evolution of geometry and properties with continued loading. Fourth, the validation of the large displacement process is not addressed in this study. Future research should focus on exploring this aspect.

Additionally, the values of physics parameters, such as friction and restitution coefficients, play a pivotal role in dynamics simulations within physics engines. Our experiments observed that these parameters produce nonlinearity in the PBR responses to ground motions. For example, we found several thresholds for friction coefficients. Within these thresholds, the friction coefficient displays various nonlinear properties. Generally, a significantly high friction coefficient made PBRs more fragile compared to a very low coefficient. However, within certain threshold ranges, the influence of varying friction is less pronounced. To quantitatively measure this nonlinearity, we recognize the necessity for additional experiments in our future research endeavors.

Future work could bypass Gazebo and directly build a VSR in the Bullet physics engine. Gazebo simplified the simulation configuration and provided perception and control packages compatible with ROS. However, when passing configuration parameters to the Bullet physics engine, Gazebo reduced the number precision of some parameters, such as the bits for floating point numbers. This reduction in the number precision presents challenges in calibrating contact properties. Building a VSR directly in Bullet allows complete control of the configuration parameters and aids in contact physics calibration. Additionally, as a previous study has used Bullet to simulate the crushing process of granular materials (Zhu and Zhao, 2019), modeling PBR simulations in Bullet allows for the study of how overturned rocks are crushed along a trajectory, providing insights into more complex modeling of rocky slope development. Simulating the crushing process also enables the examination of the impact of PBR deformation on fragility (Saifullah and Wittich, 2021). Generally, physics engines hold promise in various physics-based scenario simulations, including testing of dynamic rupture model outputs (e.g., Lozos et al., 2015).

Future work should use the VSR to build ground motion models for PSHA. For the Double Rock site, for example, we can examine the effects of the surrounding pedestals and ground motion directions on hazard curves. This study has demonstrated a forward model of large displacement dynamics, which are useful for rockfall prediction and ground motion study. At the same time, we are conceptualizing the idea of using the tool to trace the origin of overturned rocks. For example, we can simulate a large number of trajectories from various initial conditions to build a probability heatmap of the origin of the overturned rock denoted in the blue box in Fig. 15a. Using the VSR, we can also simulate a large number of PBRs with various shapes and dimensions to study the effects of ground motions on PBR distributions.

Combined with our research of using robots and machine learning for rock detection and mapping, the VSR presents a paradigm of rock detection-mapping-analysis for automated geoscience. In the future, the VSR could be installed on a companion computer of a UAV that is also developed using ROS. Once the UAV detects and maps a PBR, the VSR can rapidly analyze the PBR fragility, facilitating field data collection.

7 Conclusion

The development of the VSR has demonstrated the potential of using robotics and physics engines for studying PBR dynamics. The advances in simulating PBR overturning and large-displacement processes by the VSR provide valuable information for seismic hazard analysis, PBR fate study, rocky slope development, and rockfall prediction. Validation experiments confirmed the good performance of the velocity controller in the VSR. To validate the overturning dynamics, we compared the overturning results from the VSR with those from previous experiments and the mini shake robot. The VSR produced consistent results with previous 2D studies in the cuboid overturning experiments. The overturning response diagrams suggested that ground motion directions had complex effects on PBR fragility. We investigated the effects of surrounding pedestals on overturning responses, which had been under-explored in previous studies. The effects of the surrounding pedestals generally reduced the PBR fragility compared with flat terrains and varied with ground motion directions. For the study of the large-displacement process, we conducted 2500 experiments with different ground motions and plotted 692 trajectories where the PBR was toppled. Correlation analysis showed that the ground motions positively correlated with large displacement statistics such as mean trajectory length, mean largest velocity, and mean largest terminal distance. The rapid deployment of the VSR facilitates qualitative analysis of PBR fragility, aiding in field data collection and assessment. As a result, the VSR provides a screening method to identify the most fragile PBRs in the field for more detailed dynamics analysis. Overall, the VSR represents a significant step forward in studying PBR dynamics, providing valuable insights for researchers and practitioners.

Acknowledgements

This research was supported by the Southern California Earthquake Center (Contribution #12703), the Pacific Gas and Electric Company, and National Science Foundation award CNS-1521617. Thank you to Khalid Saifullah for providing information about DEM implementation details and computational time. We would like to extend our gratitude to the editor, Dr. Pablo Heresi, and the two anonymous reviewers for their valuable feedback and constructive critiques, which have greatly improved the quality of this work.

Data and Code Availability

The Virtual Shake Robot Software for this study is available on Zenodo (Chen, 2023). The Zenodo repository also includes experiment videos. The latest version is maintained on Github: https://github.com/DREAMS-lab/pbr_gazebo.

Competing interests

The authors declare no competing interests.

References

Agliardi, F. and Crosta, G. High resolution three-dimensional numerical modelling of rockfalls. *International Journal of Rock Mechanics and Mining Sciences*, 40(4):455–471, June 2003. doi: 10.1016/s1365-1609(03)00021-2.

Alian, M., Ein-Mozaffari, F., and Upreti, S. R. Analysis of the mixing of solid particles in a plowshare mixer via discrete element method (DEM). *Powder Technology*, 274:77–87, Apr. 2015. doi: 10.1016/j.powtec.2015.01.012.

Anderson, J. G., Biasi, G. P., and Brune, J. N. *Chapter 14 - Precarious Rocks: Providing Upper Limits on Past Ground Shaking from Earthquakes*, page 377–403. Elsevier, 2014. doi: 10.1016/b978-0-12-394848-9.00014-6.

Anooshehpoor, A., Brune, J. N., and Zeng, Y. Methodology for obtaining constraints on ground motion from precariously balanced rocks. *Bulletin of the Seismological Society of America*, 94 (1):285–303, 2004.

Asaf, Z., Rubinstein, D., and Shmulevich, I. Determination of discrete element model parameters required for soil tillage. *Soil and Tillage Research*, 92(1–2):227–242, Jan. 2007. doi: 10.1016/j.still.2006.03.006.

Azzoni, A., La Barbera, G., and Zaninetti, A. Analysis and prediction of rockfalls using a mathematical model. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 32(7):709–724, Oct. 1995. doi: 10.1016/0148-9062(95)00018-c.

Baker, J., Bradley, B., and Stafford, P. *Seismic Hazard and Risk Analysis*. Cambridge University Press, Oct. 2021. doi: 10.1017/9781108425056.

Bao, Y. and Konstantinidis, D. Dynamics of a sliding-rocking block considering impact with an adjacent wall. *Earthquake Engineering & Structural Dynamics*, 49(5):498–523, Feb. 2020. doi: 10.1002/eqe.3250.

Bell, J. W., Brune, J. N., Liu, T., Zreda, M., and Yount, J. C. Dating precariously balanced rocks in seismically active parts of California and Nevada. *Geology*, 26(6):495, 1998. doi: 10.1130/0091-7613(1998)026<0495:dpbris>2.3.co;2.

Bender, J., Erleben, K., and Trinkle, J. Interactive Simulation of Rigid Body Dynamics in Computer Graphics. *Computer Graphics Forum*, 33(1):246–270, Dec. 2013. doi: 10.1111/cgf.12272.

Beverloo, W., Leniger, H., and van de Velde, J. The flow of granular solids through orifices. *Chemical Engineering Science*, 15(3–4): 260–269, Sept. 1961. doi: 10.1016/0009-2509(61)85030-6.

Bozzolo, D. and Pamini, R. Simulation of rock falls down a valley side. *Acta Mechanica*, 63(1–4):113–130, Nov. 1986. doi: 10.1007/bf01182543.

Brune, J. N. Precariously balanced rocks and ground-motion maps for Southern California. *Bulletin of the Seismological Society of America*, 86(1A):43–54, Feb. 1996. doi: 10.1785/bssa08601a0043.

Brune, J. N., Anooshehpoor, A., Purvance, M. D., and Brune, R. J. Band of precariously balanced rocks between the Elsinore and San Jacinto, California, fault zones: Constraints on ground motion for large earthquakes. *Geology*, 34(3):137, 2006. doi: 10.1130/g22127.1.

Byerlee, J. Friction of Rocks. In *Rock Friction and Earthquake Prediction*, page 615–626. Birkhäuser Basel, 1978. doi: 10.1007/978-3-0348-7182-2_4.

Çaktı, E., Saygılı, Ö., Lemos, J. V., and Oliveira, C. S. Discrete element modeling of a scaled masonry structure and its validation. *Engineering Structures*, 126:224–236, Nov. 2016. doi: 10.1016/j.engstruct.2016.07.044.

Catanoso, D., Stucky, T., Case, J., and Rogg, A. Analysis of Sample Acquisition Dynamics Using Discrete Element Method. In *2020 IEEE Aerospace Conference*. IEEE, Mar. 2020. doi: 10.1109/aero47225.2020.9172431.

Caviezel, A., Lu, G., Demmel, S., Ringenbach, A., Bühler, Y., Christen, M., and Bartelt, P. RAMMS:: ROCKFALL-a modern 3-dimensional simulation tool calibrated on real world data. In *53rd US Rock Mechanics/Geomechanics Symposium*. OnePetro, 2019.

Chen, Z. *Automated Geoscience with Robotics and Machine Learning: A New Hammer of Rock Detection, Mapping, and Dynamics Analysis*. Phd dissertation, Arizona State University, 2022.

Chen, Z. Virtual Shake Robot Software, 2023. doi: 10.5281/ZENODO.8422820.

Chen, Z., Keating, D., Shethwala, Y., Saravanakumaran, A. A. P., Arrowsmith, R., Madugo, C., Kottke, A., and Das, J. Shakebot: A Low-cost, Open-source Shake Table for Ground Motion Seismic Studies. *arXiv preprint arXiv:2212.10763*, 2022.

Chen, Z., Arrowsmith, J. R., and Das, J. Target-oriented UAV Mapping of Precariously Balanced Rocks for Seismic Studies. *Journal of Field Robotics*, 2024, in prep.

Chiama, K., Chauvin, B., Plesch, A., Moss, R., and Shaw, J. H. Geomechanical Modeling of Ground Surface Deformation Associated with Thrust and Reverse-Fault Earthquakes: A Distinct Element Approach. *Bulletin of the Seismological Society of America*, 113(4):1702–1723, Apr. 2023. doi: 10.1785/0120220264.

Chuhan, Z., Pekau, O., Feng, J., and Guanglun, W. Application of distinct element method in dynamic analysis of high rock slopes and blocky structures. *Soil Dynamics and Earthquake Engineering*, 16(6):385–394, Aug. 1997. doi: 10.1016/s0267-7261(97)00012-2.

Cleary, P. W. A multiscale method for including fine particle effects in DEM models of grinding mills. *Minerals Engineering*, 84:88–99, Dec. 2015. doi: 10.1016/j.mineng.2015.10.008.

Coetzee, C. and Lombard, S. Discrete element method modelling of a centrifugal fertiliser spreader. *Biosystems Engineering*, 109(4):308–325, Aug. 2011. doi: 10.1016/j.biosystemseng.2011.04.011.

Coumans, E. and Bai, Y. PyBullet, a Python module for physics simulation for games, robotics and machine learning. http://pybullet.org, 2016.

Cundall, P. Formulation of a three-dimensional distinct element model—Part I. A scheme to detect and represent contacts in a system composed of many polyhedral blocks. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 25(3):107–116, June 1988. doi: 10.1016/0148-9062(88)92293-0.

Cundall, P. A. and Strack, O. D. L. A discrete numerical model for granular assemblies. *Géotechnique*, 29(1):47–65, Mar. 1979. doi: 10.1680/geot.1979.29.1.47.

Dorren, L. K. A review of rockfall mechanics and modelling approaches. *Progress in Physical Geography: Earth and Environment*, 27(1):69–87, Mar. 2003. doi: 10.1191/0309133303pp359ra.

Drumwright, E., Hsu, J., Koenig, N., and Shell, D. *Extending Open Dynamics Engine for Robotics Simulation*, page 38–50. Springer Berlin Heidelberg, 2010. doi: 10.1007/978-3-642-17319-6_7.

Erez, T., Tassa, Y., and Todorov, E. Simulation tools for model-based robotics: Comparison of Bullet, Havok, MuJoCo, ODE and PhysX. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2015. doi: 10.1109/icra.2015.7139807.

Erleben, K., Sporring, J., Henriksen, K., and Dohlmann, H. *Physics-based animation*. Charles River Media Hingham, 2005.

Garcia, F. E. and Bray, J. D. Distinct element simulations of earthquake fault rupture through materials of varying density. *Soils and Foundations*, 58(4):986–1000, Aug. 2018. doi: 10.1016/j.sandf.2018.05.009.

Garcia, F. E. and Bray, J. D. Discrete element analysis of earthquake surface fault rupture through layered media. *Soil Dynamics and Earthquake Engineering*, 152:107021, Jan. 2022. doi: 10.1016/j.soildyn.2021.107021.

Garcia-Hernandez, A., Wan, L., Dopazo-Hilario, S., Chiarelli, A., and Dawson, A. Generation of virtual asphalt concrete in a physics engine. *Construction and Building Materials*, 286:122972, June 2021. doi: 10.1016/j.conbuildmat.2021.122972.

Gascuel, J.-D., Cani-Gascuel, M.-P., Desbrun, M., Leroi, E., and Mirgon, C. *Simulating Landslides for Natural Disaster Prevention*, page 1–12. Springer Vienna, 1999. doi: 10.1007/978-3-7091-6375-7_1.

Guzzetti, F., Crosta, G., Detti, R., and Agliardi, F. STONE: a computer program for the three-dimensional simulation of rockfalls. *Computers & Geosciences*, 28(9):1079–1093, Nov. 2002. doi: 10.1016/s0098-3004(02)00025-0.

Haddad, D. E., Zielke, O., Arrowsmith, J. R., Purvance, M. D., Haddad, A. G., and Landgraf, A. Estimating two-dimensional static stabilities and geomorphic settings of precariously balanced rocks from unconstrained digital photographs. *Geosphere*, 8(5):1042–1053, 2012.

Hao, W., Jian, H., Xiang, H., Jingqing, Y., Zicheng, H., and Ying, H. A New Rockfall Simulation Tool based on Unity 3D. *IOP Conference Series: Earth and Environmental Science*, 861(3):032059, Oct. 2021. doi: 10.1088/1755-1315/861/3/032059.

Haron, A. and Ismail, K. A. Coefficient of restitution of sports balls: A normal drop test. *IOP Conference Series: Materials Science and Engineering*, 36:012038, Sept. 2012. doi: 10.1088/1757-899x/36/1/012038.

He, H., Zheng, J., and Li, Z. Comparing Realistic Particle Simulation Using Discrete Element Method and Physics Engine. In *Geo-Congress 2020*. American Society of Civil Engineers, Feb. 2020. doi: 10.1061/9780784482803.050.

He, H., Zheng, J., Chen, Y., and Ning, Y. Physics engine based simulation of shear behavior of granular soils using hard and soft contact models. *Journal of Computational Science*, 56:101504, Nov. 2021. doi: 10.1016/j.jocs.2021.101504.

Housner, G. W. The behavior of inverted pendulum structures during earthquakes. *Bulletin of the Seismological Society of America*, 53(2):403–417, Feb. 1963. doi: 10.1785/bssa0530020403.

Ishiyama, Y. Motions of rigid bodies and criteria for overturning by earthquake excitations. *Earthquake Engineering & Structural Dynamics*, 10(5):635–650, Sept. 1982. doi: 10.1002/eqe.4290100502.

Itasca Consulting Group, Inc. 3DEC — Three-Dimensional Distinct Element Code. https://www.itascacg.com/software/3dec, 2020.

Izadi, E. and Bezuijen, A. Simulation of granular soil behaviour using the Bullet physics library. In *Geomechanics from Micro to Macro*, page 1565–1570. CRC Press, Aug. 2014. doi: 10.1201/b17395-285.

Kim, J.-C., Jung, H., Kim, S., and Chung, K. Slope Based Intelligent 3D Disaster Simulation Using Physics Engine. *Wireless Personal Communications*, 86(1):183–199, June 2015. doi: 10.1007/s11277-015-2788-1.

Kirkpatrick, P. Seismic measurements by the overthrow of columns. *Bulletin of the Seismological Society of America*, 17(2):95–109, June 1927. doi: 10.1785/bssa0170020095.

Kobayashi, Y., Harp, E. L., and Kagawa, T. Simulation of rockfalls triggered by earthquakes. *Rock Mechanics and Rock Engineering*, 23(1):1–20, 1990. doi: 10.1007/bf01020418.

Koenig, N. and Howard, A. Design and use paradigms for gazebo, an open-source multi-robot simulator. In *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566)*, IROS-04. IEEE. doi: 10.1109/iros.2004.1389727.

Komaragiri, S., Gigliotti, A., and Bhasin, A. Calibration and extended validation of a virtual asphalt mixture compaction model using bullet physics engine. *Construction and Building Materials*, 311:125257, Dec. 2021. doi: 10.1016/j.conbuildmat.2021.125257.

Konstantinidis, D. A. *Experimental and analytical studies on the seismic response of freestanding and anchored building contents*. Phd dissertation, University of California, Berkeley, 2008.

Lan, H., Derek Martin, C., and Lim, C. RockFall analyst: A GIS extension for three-dimensional and spatially distributed rockfall hazard modeling. *Computers & Geosciences*, 33(2):262–279, Feb. 2007. doi: 10.1016/j.cageo.2006.05.013.

Leine, R. I., Schweizer, A., Christen, M., Glover, J., Bartelt, P., and Gerber, W. Simulation of rockfall trajectories with consideration of rock shape. *Multibody System Dynamics*, 32(2):241–271, Sept. 2013. doi: 10.1007/s11044-013-9393-4.

Lozos, J. C., Harris, R. A., Murray, J. R., and Lienkaemper, J. J. Dynamic rupture models of earthquakes on the Bartlett Springs Fault, Northern California. *Geophysical Research Letters*, 42(11): 4343–4349, June 2015. doi: 10.1002/2015gl063802.

Lu, G., Third, J., and Müller, C. Discrete element models for non-spherical particle systems: From theoretical developments to applications. *Chemical Engineering Science*, 127:425–465, May 2015. doi: 10.1016/j.ces.2014.11.050.

Ma, Q., Parshttam, S., and Montalla, M. Modelling rocking behaviour using physics engine simulation. In *Eleventh U.S. National Conference on Earthquake Engineering*, 2018.

Martin, C., Bouvard, D., and Shima, S. Study of particle rearrangement during powder compaction by the Discrete Element Method. *Journal of the Mechanics and Physics of Solids*, 51(4): 667–693, Apr. 2003. doi: 10.1016/s0022-5096(02)00101-1.

Mehrotra, A. and DeJong, M. The Performance of Slender Monuments during the 2015 Gorkha, Nepal, Earthquake. *Earthquake Spectra*, 33:321–343, Dec. 2017. doi: 10.1193/120616eqs223m.

Millington, I. Game Physics Engine Development, Mar. 2007. doi: 10.1201/9781482267327.

Milne, J. and Omori, F. On the overturning and fracturing of Brick and other Columns by horizontally applied Motion. *Seismological Journal of Japan*, 17(2):95–109, 1893.

Mirtich, B. and Canny, J. Impulse-based simulation of rigid bodies. In *Proceedings of the 1995 symposium on Interactive 3D graphics - SI3D '95*, SI3D '95. ACM Press, 1995. doi: 10.1145/199404.199436.

Papantonopoulos, C., Psycharis, I. N., Papastamatiou, D. Y., Lemos, J. V., and Mouzakis, H. P. Numerical prediction of the earthquake response of classical columns using the distinct element method. *Earthquake Engineering & Structural Dynamics*, 31(9): 1699–1717, June 2002. doi: 10.1002/eqe.185.

Pezo, L., Jovanović, A., Pezo, M., Čolović, R., and Lončar, B. Modified screw conveyor-mixers – Discrete element modeling approach. *Advanced Powder Technology*, 26(5):1391–1399, Sept. 2015. doi: 10.1016/j.apt.2015.07.016.

Pompei, A., Scalia, A., and Sumbatyan, M. A. Dynamics of Rigid Block due to Horizontal Ground Motion. *Journal of Engineering Mechanics*, 124(7):713–717, July 1998. doi: 10.1061/(asce)0733-9399(1998)124:7(713).

Purvance, M. D. *Overturning of slender blocks: Numerical investigation and application to precariously balanced rocks in southern California*. Phd dissertation, University of Nevada, Reno, 2005.

Purvance, M. D., Anooshehpoor, A., and Brune, J. N. Freestanding block overturning fragilities: Numerical simulation and experimental validation. *Earthquake Engineering & Structural Dynamics*, 37(5):791–808, Jan. 2008. doi: 10.1002/eqe.789.

Purvance, M. D., Anooshehpoor, A., and Brune, J. N. Fragilities for Precarious Rocks at Yucca Mountain. techreport, Pacific Earthquake Engineering Research Center, 2012.

Pytlos, M., Gilbert, M., and Smith, C. C. Modelling granular soil behaviour using a physics engine. *Géotechnique Letters*, 5(4): 243–249, Dec. 2015. doi: 10.1680/jgele.15.00067.

Rood, A. H., Rood, D. H., Stirling, M. W., Madugo, C. M., Abrahamson, N. A., Wilcken, K. M., Gonzalez, T., Kottke, A., Whittaker, A. C., Page, W. D., and Stafford, P. J. Earthquake Hazard Uncertainties Improved Using Precariously Balanced Rocks. *AGU Advances*, 1 (4), Oct. 2020. doi: 10.1029/2020av000182.

Rood, A. H., Rood, D. H., Balco, G., Stafford, P. J., Ludwig, L. G., Kendrick, K. J., and Wilcken, K. M. Validation of earthquake ground-motion models in southern California, USA, using precariously balanced rocks. *GSA Bulletin*, Dec. 2022. doi: 10.1130/b36484.1.

Saifullah, M. and Wittich, C. Fragility of Precariously Balanced Rocks: Shake Table Testing and Numerical Modeling for a Sample Granitic Rock. In *12th National Conference on Earthquake Engineering*, 2022.

Saifullah, M. K. and Wittich, C. E. Seismic Response of Two Freestanding Statue-Pedestal Systems during the 2014 South Napa Earthquake. *Journal of Earthquake Engineering*, 26(10): 5086–5108, Jan. 2021. doi: 10.1080/13632469.2020.1859004.

Scalia, A. and Sumbatyan, M. A. Slide Rotation of Rigid Bodies Subjected to a Horizontal Ground Motion. *Earthquake Engineering & Structural Dynamics*, 25(10):1139–1149, Oct. 1996. doi: 10.1002/(sici)1096-9845(199610)25:10<1139::aid-eqe606>3.0.co;2-s.

SCEC researchers. SCEC Precariously Balanced Rocks Project: PBR Data Archive, 2022. https://www.scec.org/project/pbr.

Shenton, I. and Harry, W. Criteria for Initiation of Slide, Rock, and Slide-Rock Rigid-Body Modes. *Journal of Engineering Mechanics*, 122(7):690–693, July 1996. doi: 10.1061/(asce)0733-9399(1996)122:7(690).

Shi, B., Anooshehpoor, A., Zeng, Y., and Brune, J. N. Rocking and overturning of precariously balanced rocks by earthquakes. *Bulletin of the Seismological Society of America*, 86(5): 1364–1371, Oct. 1996. doi: 10.1785/bssa0860051364.

Stanford Artificial Intelligence Laboratory. Robotic Operating System, 2018. https://www.ros.org.

Sun, C., Nakashima, H., Shimizu, H., Miyasaka, J., and Ohdoi, K. Physics engine application to overturning dynamics analysis on banks and uniform slopes for an agricultural tractor with a rollover protective structure. *Biosystems Engineering*, 185:150–160, Sept. 2019. doi: 10.1016/j.biosystemseng.2018.06.005.

Toson, P. and Khinast, J. G. Impulse-based dynamics for studying quasi-static granular flows: Application to hopper emptying of non-spherical particles. *Powder Technology*, 313:353–360, May 2017. doi: 10.1016/j.powtec.2017.03.016.

van den Bergen, G. *Collision Detection in Interactive 3D Environments*. CRC Press, Oct. 2003. doi: 10.1201/9781482297997.

Veeraraghavan, S. *Toppling analysis of precariously balanced rocks*. Phd dissertation, California Institute of Technology, 2015.

Veeraraghavan, S., Hudnut, K. W., and Krishnan, S. Toppling Analysis of the Echo Cliffs Precariously Balanced Rock. *Bulletin of the Seismological Society of America*, 107(1):72–84, Dec. 2016. doi: 10.1785/0120160169.

Veeraraghavan, S., Hall, J. F., and Krishnan, S. Modeling the Rocking and Sliding of Free-Standing Objects Using Rigid Body Dynamics. *Journal of Engineering Mechanics*, 146(6), June 2020. doi: 10.1061/(asce)em.1943-7889.0001739.

Williams, J., Lu, Y., and Trinkle, J. C. A Complementarity Based Contact Model for Geometrically Accurate Treatment of Polytopes in Simulation. In *Volume 6: 10th International Conference on Multibody Systems, Nonlinear Dynamics, and Control*, IDETC-CIE2014. American Society of Mechanical Engineers, Aug. 2014. doi: 10.1115/detc2014-35231.

Wittich, C. and Hutchinson, T. Rocking bodies with arbitrary interface defects: Analytical development and experimental verification. *Earthquake Engineering & Structural Dynamics*, 47(1): 69–85, July 2017. doi: 10.1002/eqe.2937.

Xu, Z., Lu, X., Guan, H., Han, B., and Ren, A. Seismic damage simulation in urban areas based on a high-fidelity structural model and a physics engine. *Natural Hazards*, 71(3):1679–1693, Dec. 2013. doi: 10.1007/s11069-013-0972-8.

Yim, C., Chopra, A. K., and Penzien, J. Rocking response of rigid blocks to earthquakes. *Earthquake Engineering & Structural Dy-*

namics, 8(6):565–587, Jan. 1980. doi: 10.1002/eqe.4290080606.

Zhu, F. and Zhao, J. Modeling continuous grain crushing in granular media: A hybrid peridynamics and physics engine approach. *Computer Methods in Applied Mechanics and Engineering*, 348: 334–355, May 2019. doi: 10.1016/j.cma.2019.01.017.

doi:10.26443/seismica.v3i1.624

SEISMICA

Dispersive Elastic Moduli and Frequency-Dependent Attenuation due to Wave-Induced Fluid Flow in Metapelite

C. Fliedner * [1], **M. E. French** [1]

[1]Department of Earth, Environmental and Planetary Sciences, Rice University, Houston, United States of America

Author contributions: *Conceptualization*: CF, MF. *Methodology*: CF. *Software*: CF. *Validation*: CF. *Formal Analysis*: CF. *Investigation*: CF. *Writing - original draft*: CF, MF. *Writing - Review & Editing*: CF, MF. *Visualization*: CF. *Supervision*: MF. *Funding acquisition*: CF, MF.

Abstract Seismic waves are used to interpret geologic structure, composition, and environmental conditions in the Earth. However, rocks are not perfectly elastic and their viscoelasticity dissipates energy during wave propagation. In saturated rocks, wave-induced fluid flow mechanisms can cause viscoelasticity resulting in frequency-dependent attenuation, velocities, and elastic moduli (dispersion). In subduction zones, some regions exhibit evidence of overpressurized fluids where dispersion and attenuation are hypothesized to be important in interpreting fault slip behavior from seismic waves. However, their importance has not been well-characterized because of a lack of measurements on relevant lithologies and under saturated conditions. We measured the Young's and shear moduli and the attenuation of a greenschist facies metapelite with the forced oscillation technique at frequencies between 2×10^{-5} and 30 Hz. The moduli and attenuation are frequency-dependent under saturated conditions and depend on effective pressure. At relatively low effective pressure, the Young's and shear moduli increase by over 50% between 2×10^{-5} and 30 Hz. We use Standard Linear Solid viscoelastic models to investigate the relationship between the attenuation and dispersion in the Orocopia schist. The models agree with the experimental data and demonstrate that viscoelasticity causes significant dispersion and attenuation in subduction zones, affecting our interpretation of earthquakes.

Production Editor:
Garreth Funning
Handling Editor:
Pathikrit Bhattacharya
Copy & Layout Editor:
Ethan Williams

Signed reviewer(s):
David Healy, Nir Badt

Received:
27 March 2023
Accepted:
26 November 2023
Published:
26 January 2024

Non-technical summary Seismic waves from earthquakes are used to image and interpret the composition, structure, and environmental conditions of the Earth. Since rocks are not perfectly elastic, waves slow down and lose energy during propagation. The decrease in wave energy (i.e., attenuation) manifests as an amplitude reduction and velocity changes reflect changes in elastic moduli. Pore fluids cause both attenuation and changes in elastic moduli. To interpret fault slip with seismic waves, we must understand how they evolve as they reach the surface. Measurements of attenuation and elasticity relevant to fluid-rich areas in active subduction zones are limited but crucial. We conducted laboratory experiments on schist, common in subduction zones, to measure its attenuation and elasticity using the 'forced oscillation' method. Hence, we apply a cyclical force at a specific frequency, mimicking wave frequencies of earthquakes, and measure the induced deformation. We analyzed the amplitude and time difference between the force and the deformation to determine the rock properties. The properties vary with frequency and increase by over 50% when fluid content is high. Using physics models, we conclude that rock properties change because of fluid movement in the pores. Consequently, fluids significantly affect how we interpret rock properties in subduction zones.

1 Introduction

Elastic waves are one of the most powerful tools for constraining processes in Earth's interior. They are used to image Earth's structure, interpret its composition and environmental conditions such as temperature and fluids, and constrain the physics of processes like fault slip. The amplitudes of elastic waves decay as they propagate away from their source (Anderson and Archambeau, 1964; Karato, 1993). This energy dissipation is called attenuation which results from viscoelastic deformation of the rock and is quantified using the quality factor Q (Anderson and Archambeau, 1964; Brennan and Stacey, 1977). In Earth's upper crust, attenuation $(1/Q)$, occurs primarily due to fluid flow between pores and microcracks, which dissipates energy (O'Connell and Budiansky, 1974, 1977; Winkler and Nur, 1979; Bernabé and Revil, 1995; Borgomano et al., 2019). As a result of this energy dissipation, saturated rocks attenuate elastic waves and the elastic moduli and wave velocities depend on frequency, which is called dispersion (O'Connell and Budiansky, 1977; Winkler and Nur, 1979; Spencer, 1981). To make geologic interpretations on the basis of seismic velocities and attenuation, it is crucial to determine how they depend on wave frequency and to extrapolate them across geologic conditions requires determining the underlying processes that control attenuation and dispersion.

Because of the effects of water on wave propagation, attenuation and dispersion of elastic moduli may be significant in regions of subduction zones with high pore

*Corresponding author: celine.fliedner@rice.edu

pressure, although the rocks and conditions are not as well studied as those of the upper crust. Regions of anomalously low P-wave (V_P) and S-wave (V_S) velocities and high V_P/V_S are sometimes observed below the seismogenic zone of megathrusts where diverse modes of fault slip, including tremors, low-frequency earthquakes (LFEs), and very low-frequency earthquakes (VLFEs) are also observed (Audet et al., 2009; Shelly et al., 2006; Bostock et al., 2012; Delph et al., 2018; Calvert et al., 2020). Field observations reveal the presence of low-porosity metapelite rocks rich in phyllosilicates, as well as extensive veining which supports the geophysical evidence of high pore pressure (Philippot and Selverstone, 1991; Angiboust et al., 2015; Muñoz-Montecinos et al., 2021; Condit and French, 2022). High pore pressure at these depths is thought to be caused by the release of water during dehydration reactions during greenschist to blueschist facies metamorphism (Peacock, 1987; Muñoz-Montecinos et al., 2021; Tewksbury-Christle et al., 2021; Condit et al., 2022). The dehydration of metabasalt and metapelite in particular, may release enough fluid to create high pore pressure at depths of slow slip and low-frequency events (Condit et al., 2020). However, although seismic imaging provides information that we use to infer fluid conditions at depth, we know little about how seismic waves are altered as they propagate through the rocks at these conditions.

Correlations between diverse modes of fault slip and evidence for high fluid pressure have resulted in the prominent hypothesis that fluid pressure somehow causes tremors, LFEs, and/or VLFEs. An alternative hypothesis is that attenuation and dispersion caused by high pore fluid pressure alter the seismic waves of small typical earthquakes such that they appear as LFEs and/or VLFEs by the time the waves are recorded at the surface (Gomberg et al., 2012; Bostock et al., 2017; Littel et al., 2018; Nakai et al., 2021). Specifically, the frequency range of typical earthquakes ranges from 1 to 30 Hz, but LFEs and VLFEs are depleted in high frequencies and primarily exhibit wave frequencies between 0.1 and 8 Hz (Farge et al., 2020; Supino et al., 2020; Ide et al., 2007; Obara, 2002; Obara and Kato, 2016; Shelly et al., 2007; Thomas et al., 2016; Bostock et al., 2015). It is unclear if the low frequencies of these events are caused by attenuation of the waves or a slower rupture and slip mechanism at the source (Ito et al., 2007; Gomberg et al., 2016; Shapiro et al., 2018; Sammis and Bostock, 2021; Wei et al., 2021). We previously showed that overpressurized fluids can cause attenuation in the laboratory that is sufficient to deplete high frequencies during seismic wave propagation (Fliedner and French, 2023j). Here we use additional experimental data and rock physics models to quantitatively evaluate the mechanisms of attenuation and the magnitudes of velocity dispersion, and then extrapolate the results to in-situ conditions where LFEs and VLFEs occur beneath the seismogenic zone.

Previous studies demonstrate that the elastic moduli of porous sedimentary rocks are dispersive at seismic frequencies (0.1 - 30 Hz) because of wave-induced fluid flow occurring in the pore space (Farge et al., 2020; Ide et al., 2007; Obara, 2002; Obara and Kato, 2016; Thomas et al., 2016; Shelly et al., 2007; Supino et al., 2020). However, measurements of dispersion and attenuation on the lithologies present near the base of the subduction seismogenic zone are scarce (Fliedner and French, 2023j). The thin elongate pores characteristic of schists are particularly compressible under load (Walsh, 1965; Mavko and Nur, 1978; Kranz, 1983) and the small stress perturbations caused by propagating waves can close the pores and cause wave-induced fluid flow (White, 1975; Mavko and Nur, 1975; Pride et al., 2004), resulting in considerable attenuation (Toksöz et al., 1979; Gomberg et al., 2012; Fliedner and French, 2021, 2023j). Wave-induced fluid flow is often observed at seismic frequencies (Mavko and Nur, 1975; Dvorkin et al., 1994; Batzle et al., 2006), because fluid has time to flow between pores. In contrast, when the wave frequency is high (ultrasonic), fluid does not have enough time to flow and the rock properties are predicted to be independent of frequency (Biot, 1956; Toksöz et al., 1979). This is one reason elastic moduli measured in the laboratory tend to differ between ultrasonic and seismic frequencies (O'Connell and Budiansky, 1977; Adelinet et al., 2010; Dvorkin et al., 1994; Spencer, 1981; Pimienta et al., 2015; Borgomano et al., 2019). The fluid flow mechanisms that cause attenuation and dispersion at seismic frequencies and their relationship to the rock microstructure remain difficult to constrain from laboratory data and are rarely reported (Fliedner and French, 2023j). Fluid substitution models can predict elastic moduli at seismic frequencies, but they have yet to be confirmed by laboratory experiments for a number of lithologies, including metapelites (Gassmann, 1951; Brown and Korringa, 1975).

We present laboratory measurements of the frequency-dependent elastic moduli of a greenschist facies metapelite, the Orocopia schist, at frequencies between 2×10^{-5} and 30 Hz under dry and saturated conditions. This study is the first to report the elastic moduli and attenuation of phyllosilicate-rich metapelite under saturated conditions and at seismic frequencies. We combine these measurements with the attenuation measurements ($1/Q$) reported in (Fliedner and French, 2023j) to demonstrate that dispersion occurs concurrently with frequency-dependent attenuation. Together, these data allow us to apply rock physics models of wave-induced fluid flow and evaluate the mechanisms responsible. Using the Standard Linear Solid viscoelastic model, we show that the measured attenuation and dispersion in elastic moduli can be explained by the occurrence of two wave-induced flow mechanisms, squirt flow and patchy saturation. With this information and an understanding of the rock microstructure, it is then possible to make predictions about attenuation at geologic conditions under different fluid conditions.

2 Orocopia Schist

We collected a sample of the Orocopia schist from the Orocopia Mountains of Southern California (Figure 1a)

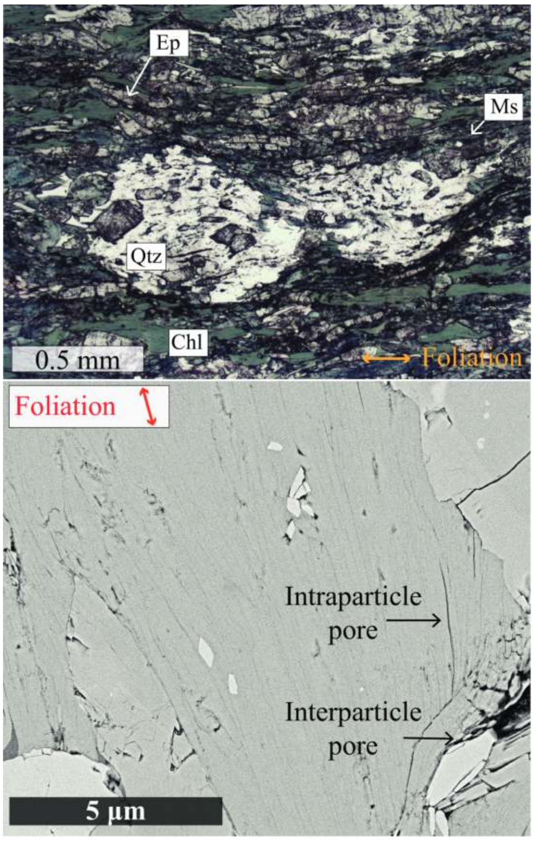


Figure 1 (a) A photomicrograph of a thin section of the Orocopia schist. Foliation is defined by aligned phyllosilicates and the orientation of individual minerals is indicated in pink. (b) Backscattered electron (BSE) images of the Orocopia schist. Foliation orientation is indicated with a red arrow. Pores are dominantly low aspect ratio (crack shaped) with the long axes parallel to the phyllosilicates and foliation plane. Figure from (Fliedner and French, 2021)

.

and the same block was used in the experiments of Fliedner and French (2021) and Fliedner and French (2023j). The Orocopia schist is a metapelite formed during Laramide subduction and reached peak conditions of 1.1 GPa pressure and 600 °C. However, it subsequently underwent greenschist facies metamorphism, which is reflected in its current mineral assemblage (Chapman et al., 2016; Jacobson and Dawson, 1995; Jacobson et al., 2007). The modal composition of our sample was reported in Fliedner and French (2021) and is 24% quartz, 31% chlorite, 22% muscovite, 17% epidote, and 3% calcite, similar to that reported in other studies of the Orocopia schist (Jacobson and Dawson, 1995; Jacobson et al., 2007; Platt et al., 2018). The schist has a well-defined foliation caused by the shape preferred orientation of phyllosilicate and quartz grains (Fliedner and French, 2021). The abundance of aligned phyllosilicates creates a primary transverse isotropic symmetry that we confirmed with ultrasonic velocity measurements and effective medium models (Fliedner and French, 2021). The sources of transverse isotropy can be seen in backscattered electron images which show phyllosilicate grains aligned within 5–10° and occasionally up to 25° to the foliation plane (Figure 1b). The schist has a secondary transverse isotropy caused by abundant thin elongated pores oriented sub-parallel to phyllosilicate grains (Figure 1). Porosity includes delamination cracks, intragranular fractures, and both elongated and equant intergranular pores. The connected porosity is approximately 1 % and was determined in the laboratory using the difference between the mass of a water-saturated core of known volume and the mass of the same core under dry conditions. The permeability normal to the foliation is 4.2×10^{-19} m^2 and was measured by CoreLaboratories using a steady state nano-permeameter under hydrostatic conditions at a pressure of 10 MPa and with nitrogen gas pore fluid at 0.2 MPa. Because our experiments were conducted with water pore fluid, not nitrogen, the water permeability may be lower due to the Klinkenberg effect which describes the slip of gases along pore walls (e.g., Tanikawa and Shimamoto (2009)). A precise correction requires a measurement of pore radii. However, previous research on rocks of similar gas permeability shows that the water permeability may be up to one order of magnitude lower, leading to a lower-bound water permeability of $\sim 4.2 \times 10^{-20}$ m^2 (Tanikawa and Shimamoto, 2009).

We measured the attenuation and elastic moduli of a single core of Orocopia schist to limit the effects of sample variability on our measurements. The core is oriented with its axis normal to foliation; this is the expected approximate orientation of wave travel away from the megathrust in subduction zones and the expected orientation of maximum attenuation, $1/Q$ (Delle Piane et al., 2014; Mikhaltsevitch et al., 2020). The experimental core has a diameter of 52.9 mm and a height of 25.4 mm. The core ends were trimmed and ground flat with a surface grinder. Prior to measurements under saturated conditions, the core was presaturated in deionized water under vacuum for seven days. In Fliedner and French (2021), we estimated the time scale of fluid diffusion in the Orocopia and determined that 12 hrs is sufficient as long as the water permeability is higher than $\sim 8 \times 10^{-22}$ m^2. Thus, 7 days should be sufficient to saturate most connected porosity, but it is possible that some gas remains trapped in disconnected pores or those connected by extremely small pore throats.

3 Methods

3.1 Experimental setup

Elastic moduli and attenuation measurements were made using a servo-controlled triaxial deformation apparatus at Rice University (Figure 2). In this apparatus, a silicone oil confining medium applies the least compressive stress (σ_3) parallel to the core radius and deionized water applies the pore fluid pressure (P_f) through the bottom end cap. The pore fluid outlet in the top end cap is closed, which prevents a dead volume of fluid that can create measurement artifacts (Dunn, 1986; Pimienta et al., 2016a). When accounting for the low porosity of the Orocopia schist and the experimental setup, dead volume is expected to have a negligible impact on our measurements (Dunn, 1986; Pimienta

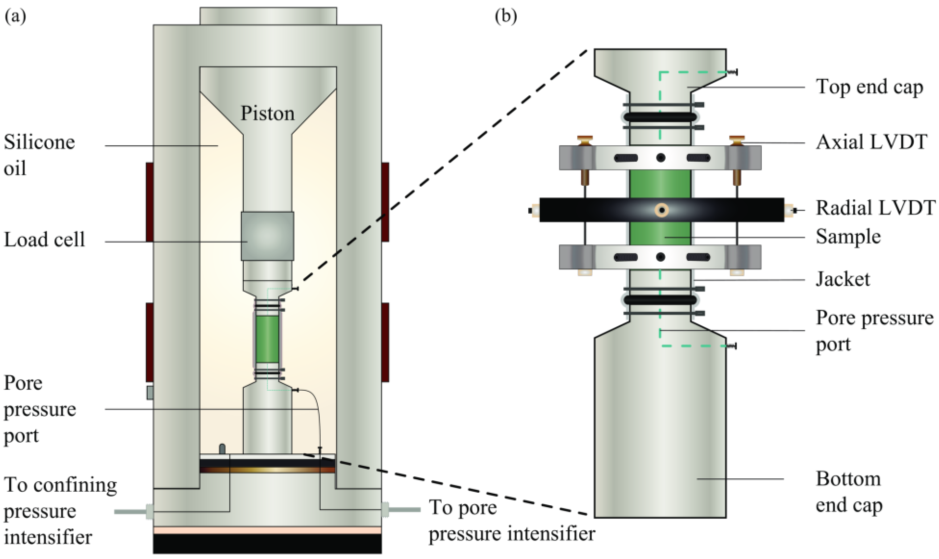


Figure 2 The triaxial deformation apparatus. (a) Diagram of the pressure vessel and (b) Magnified view of the sample configuration with the 2 axial LVDTs and 4 radial LVDTs.

et al., 2016a). The sample and pore fluid are isolated from the silicone oil with two polyolefin jackets (Figure 2b), which limits lateral fluid flow between the sample and the jackets. An axial piston applies the greatest compressive stress (σ_1) parallel to the core axis. The differential stress ($\sigma_1 - \sigma_3$) is recorded with an internal load cell having a precision of 0.3 MPa. Deformation of the sample is measured using 6 linear voltage differential transformers (LVDTs), 2 axial and 4 radial. The 2 axial LVDTs are placed at a $180°$ from one another and the 4 radial LVDTs are at a $90°$ from one another. Axial deformation was always measured with high resolution (63 nm) LVDTs and radial deformation was measured with high resolution LVDTs in 2 experiments and medium resolution (250 nm) LVDTs in 2 experiments (Table 1). The experiments were conducted at room temperature and temperature within the pressure vessel was measured. The internal temperature varied less than $1°$C.

We used the forced oscillation technique to measure the frequency-dependent attenuation and elastic moduli of the Orocopia schist and 2 well-characterized materials, aluminum alloy (Al-6061) and Poly(methyl methacrylate) (PMMA), to calibrate our measurements (Figure 1a) (Spencer, 1981; Jackson and Paterson, 1987; Batzle et al., 2006). Measurements were made at frequencies between 2×10^{-5} and 30 Hz and the attenuation measurements of Orocopia schist were previously reported in Fliedner and French (2023j). The forced oscillation technique consists of applying a small sinusoidal oscillation in stress at a discrete frequency parallel to the core axis and then measuring the induced axial and radial strains. Deformation is assumed to be anelastic, meaning completely recoverable but with some energy dissipation. Recoverable deformation is assumed because stress and strain are small, and our measurements evaluate if and how much energy is dissipated.

The Young's modulus, E is determined from the ratio of the amplitudes in stress (σ_a) and resulting axial strain (ε_a):

$$E = \frac{\sigma_a}{\varepsilon_a} \tag{1}$$

The relationship between the different elastic moduli means that the shear modulus, which is defined as the ratio of shear stress to shear strain, can be determined from the Young's modulus (Equation 1) and the Poisson's ratio, ν. The Poisson's ratio is the ratio of the radial strain (ε_r) and axial strain ($\nu = -\varepsilon_r/\varepsilon_a$). The shear modulus can be determined from the axial stress and the difference between the axial and radial strain as:

$$G = \frac{E}{2(1+\nu)} = \frac{\sigma_a}{2(\varepsilon_a - \varepsilon_r)} \tag{2}$$

The phase offset between the stress oscillation and the induced strain, ϕ, is measured in radians and gives the attenuation as $1/Q = \tan\phi$ at that frequency where Q is a property called the quality factor (Nowick and Berry, 1972). We measured the Young's modulus attenuation ($1/Q_E$) from the phase difference between the axial stress, σ_a, which has phase $\phi(\sigma_a)$, and axial strain, ε_a which has phase $\phi(\varepsilon_a)$:

$$1/Q_E = \tan\big(\phi(\sigma_a) - \phi(\varepsilon_a)\big) \tag{3}$$

The shear modulus attenuation (Q_S) is given by the phase difference between axial stress, σ_a, and the difference between axial strain (ε_a with phase $\phi(\varepsilon_a)$) and radial strain (ε_r with phase $\phi(\varepsilon_r)$) (Yin et al., 2019):

$$1/Q_S = \tan\Big(\phi(\sigma_a) - \phi(\varepsilon_a - \varepsilon_r)\Big) \tag{4}$$

We conducted two experiments on an aluminum core (G0188, G0198) and two experiments on PMMA (G0184, G0197) as summarized in Table 1 (Fliedner and French, 2023a,f,c,g). Core dimensions for these samples are 50 mm in length and 25.4 mm in diameter. We reported the attenuation measurements for one experiment on dry Orocopia schist (G0187) and four experiments on water-saturated Orocopia schist at two different effective pressures (G0190, G0191, G0199, G0200) in Fliedner

Set	Test	Sample	σ_3	P_f	P_{eff}	$\sigma_1 - \sigma_3$	A_{pp}	N_f	Frequency range	Radial LVDT?
			MPa	*MPa*	*MPa*	*MPa*	*MPa*		*Hz*	
1	G0184	PMMA	3	0	3	3	0.4	71	1×10^{-4} - 30	No
	G0187	Orocopia	10	0	10	10	1	71	1×10^{-4} - 30	No
	G0188	Al-6061	10	0	10	10	1	69	1×10^{-4} - 30	Yes
	G0190	Orocopia	10	1	9	10	1	73	1×10^{-4} - 30	Yes
	G0191	Orocopia	10	8	2	10	1	72	1×10^{-4} - 30	Yes
2	G0197	PMMA	10	0	5	5	1	80	1×10^{-4} - 30	Yes
	G0198	Al-6061	10	0	10	10	1	13	1×10^{-4} - 10	Yes
	G0199	Orocopia	10	1	9	10	1	57	2×10^{-5} - 10	Yes
	G0200	Orocopia	10	8	2	10	1	82	2×10^{-5} - 10	Yes

Table 1 Table of the forced oscillations experiments conducted. The experiments were conducted in two separate runs to establish reproducibility, each of which is labeled as a set. The environmental conditions are the confining pressure (σ_3), the pore pressure (P_f), the effective pressure (P_{eff}), the differential stress ($\sigma_1 - \sigma_3$), and the peak-to-peak amplitude (A_{pp}) of the sinusoidal stress oscillation. N_f is the total number of measurements made over all frequencies during a given test.

and French (2023j) and show those results here along with the elastic moduli from the same experiments, which are not reported elsewhere (Fliedner and French, 2023b,d,h,e,i). All measurements were made at a confining pressure of 10 MPa except for one experiment on PMMA, which was conducted at 5 MPa (G0197). Measurements made on the Orocopia schist under saturated conditions were made at 1 and 8 MPa pore fluid pressure, resulting in Terzaghi effective pressures ($P_{eff} = \sigma_3 - P_f$) of 9 and 2 MPa (Terzaghi et al., 1996).

For measurements under dry conditions, including on aluminum and PMMA, the confining pressure was increased by 1 MPa every 5 minutes until 10 MPa and the sample equilibrated at pressure for 12 hours. For measurements under saturated conditions, an initial confining pressure of 5 MPa and a pore pressure of 0.5 MPa were applied. Next, we increased the pore pressure to 1 MPa and confining pressure to 10 MPa over 10 minutes and allowed the system to equilibrate for 12 hours (G0190 and G0199). Following measurements at these conditions, we increased the pore pressure to 8 MPa over 10 minutes and again allowed the sample to equilibrate for 12 hours (G0191 and G0200). The two separate sets of experiments were made under saturated conditions to establish the reproducibility of our measurements.

Once at experimental conditions, we applied a small differential stress of 10 MPa to assure contact between the end caps and the sample and then applied a sinusoidal stress oscillation with a peak-to-peak amplitude of 1 MPa (0.4 MPa for one experiment on PMMA), which causes axial strains of $\sim 10^{-5}$. Measurements of elastic moduli and attenuation were made at 22 discrete frequencies from 1×10^{-4} to 30 Hz on aluminum and PMMA and at 26 discrete frequencies from 2×10^{-5} to 30 Hz on the Orocopia schist. The data sampling frequency varied between 50 and 5000 Hz and increased with the frequency of the stress oscillation. Multiple measurements were made at a given frequency during an experiment and in total, we conducted between 2 and 6 forced oscillation tests at each frequency and effective stress. Experiments on the aluminum and PMMA cores we made in between the two sets of experiments made on the Orocopia schist to verify the accuracy and consistency of our measurements (Table S1 and Table S2).

3.2 Calibration

We corrected the attenuation and elastic moduli to account for signal distortion which is primarily caused by electronic noise at high frequencies and changes in environmental conditions like room temperature at low frequencies. We do so using the measurements of attenuation and elastic moduli on aluminum alloy (Al-6061) and Poly(methyl methacrylate) (PMMA), which are well-characterized (Figure 3). Aluminum has very low and frequency-independent attenuation and high elastic moduli (Lakes, 2009; Duffy, 2002; Oberg and McCauley, 2020) and PMMA has a relatively high frequency-dependent attenuation and low elastic moduli (Lakes, 2009; Madonna and Tisato, 2013; Saltiel et al., 2017; Lee et al., 2000). Radial deformation was recorded during both experiments on aluminum (G0188 and G0198), but a high-resolution LVDT was only used during one of the experiments (G0188, Table 1). Radial deformation was only measured during one of two experiments on PMMA (G0197), so we do not report shear modulus or Q_S for the other (G0184).

We calibrated the elastic moduli and the attenuation measurements using equations with 4 fitted parameters, A, B, C and D (Equation 5). Each parameter accounts for different effects, with A correcting for the effect of amplifiers and B and C correcting for phase distortion effects at high and low frequencies, respectively. The coefficient D corrects for effects independent of frequency. The parameters A, B, C and D were determined as the correction required to fit our measurements on aluminum and PMMA to the published

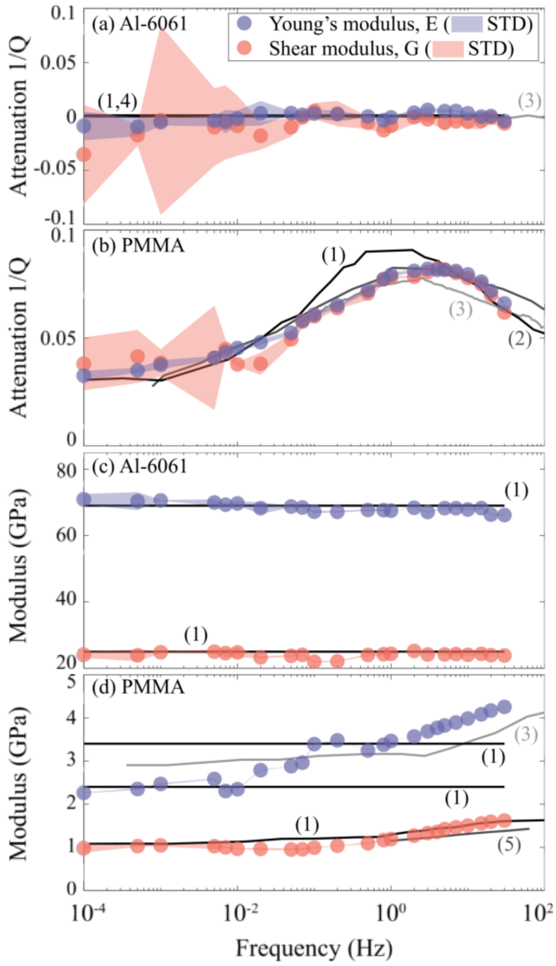


Figure 3 The results of calibration measurements on 6061 aluminum alloy (Al-6061) and Poly(methyl methacrylate) (PMMA) samples after correction using Equation 5. The Young's moduli and attenuation are shown in blue and shear moduli and attenuation are shown in pink. The blue and pink shaded areas indicate the standard deviation (STD) for the Young's and shear moduli measurements, respectively, and are extrapolated between point measurements. Published values of attenuation and elastic moduli are shown with grey lines for reference and labeled as (1) Young's and Shear moduli from Lakes (2009), (2) Shear modulus from Lee et al. (2000), (3) Young's modulus from Madonna and Tisato (2013), (4) Shear modulus from Duffy (2002), and (5) Shear modulus from Saltiel et al. (2017). (a) Attenuation $1/Q_E$ and $1/Q_S$ of Al-6061, (b) Attenuation $1/Q_E$ and $1/Q_S$ of PMMA, (c) Young's E and shear G moduli of Al-6061, and (d) Young's E and shear G moduli of PMMA.

references for these materials.

$$Data_{corrected} = A * data + B * f + C * log_{10}(f) + D \quad (5)$$

We collected calibration data during each of the two sets of measurements (Table 1). During one set an additional amplifier was used, resulting in different values of A, B, C and D for the two sets (Figures S2 to S10 in Supplementary material). For the first group (G0184, G0187, G0190, and G0191), we first fit Equation 5 to aluminum so that it conforms to published values, and then applied the same correction to the experiments on the PMMA to confirm that the corrected results are consistent with published values and then the Orocopia schist. For the second group, we fit Equation 5 to the PMMA (G0197) and the aluminum (G0198) separately to conform with published values. The Orocopia schist (G0199, G0200) was then corrected using the average parameters determined for PMMA (G0197) and Al-6061 (G0198). A table of the coefficients can be found in the Supplementary Materials (Tables S1 and Table S2).

4 Results for Elastic Moduli and Attenuation

Under water-saturated conditions, the Orocopia schist is stiffer at most frequencies and has higher attenuation at all frequencies than under dry conditions (Figure 4a and c). We also measure higher Young's and shear moduli and lower attenuation ($1/Q_E$ and $1/Q_S$) at an effective pressure of 9 MPa than at 2 MPa. In addition, whereas the elastic moduli and attenuation are independent of frequency under dry conditions, both vary with frequency under water-saturated conditions.

Under dry conditions, the Young's modulus, E, is 50 GPa (± 0.160) at all frequencies tested (Figure 4a). For comparison, the Young's modulus at ultrasonic frequencies (1×10^6 Hz), which we reported in Fliedner and French (2021), is 51 GPa indicating relatively constant stiffness over 10 orders of magnitude in frequency (Figure 4a). Under saturated conditions, the Young's modulus increases with increasing frequency (Figure 4a). For instance, at P_{eff} = 2 MPa the Young's modulus increases from 41 GPa ($\pm$ 2) to 67 GPa ($\pm$ 1) with increasing frequency from 2×10^{-5} Hz to 30 Hz, and it appears to increase in two phases (Figure 4a). At P_{eff} = 2 MPa Young's modulus increases linearly in log space from 41 to 57 GPa between 2×10^{-5} and 1×10^{-3} Hz and is then relatively constant until $\sim$ 0.1 Hz, when it again increases approximately linearly in log space. The Young's modulus increases similarly in two phases at 9 MPa effective pressure although the plateau from 1×10^{-3} to 0.1 Hz is less clear, and the modulus is consistently 3 GPa greater than at 2 MPa effective pressure. At effective pressures of 2 and 9 MPa, the Young's moduli are 5 and 8 % lower at 30 Hz than at ultrasonic frequencies (1×10^6 Hz), and this is consistent with extrapolating the frequency-dependence observed in our data an additional 5 orders of magnitude (Fliedner and French, 2021). Similar to the Young's modulus, the shear modulus is also dispersive under water-saturated conditions between 2×10^{-5} and 30 Hz and we did not measure the shear modulus under dry conditions (Figure 4b). For instance, at 2 MPa effective pressure, the shear modulus increases from 14 GPa ($\pm$ 3) to 25 GPa ($\pm$ 1) at 2×10^{-5} to 30 Hz. At a given frequency, the shear modulus measured at P_{eff} = 2 MPa effective pressure is $\sim$ 2 GPa lower than the shear modulus at 9 MPa. In contrast to the Young's modulus, the shear moduli are actually higher at 30 Hz than at 1×10^6 Hz indicating that these results cannot be directly extrapolated to ultrasonic frequencies (Fliedner and French, 2021).

Under dry conditions, attenuation of the Young's

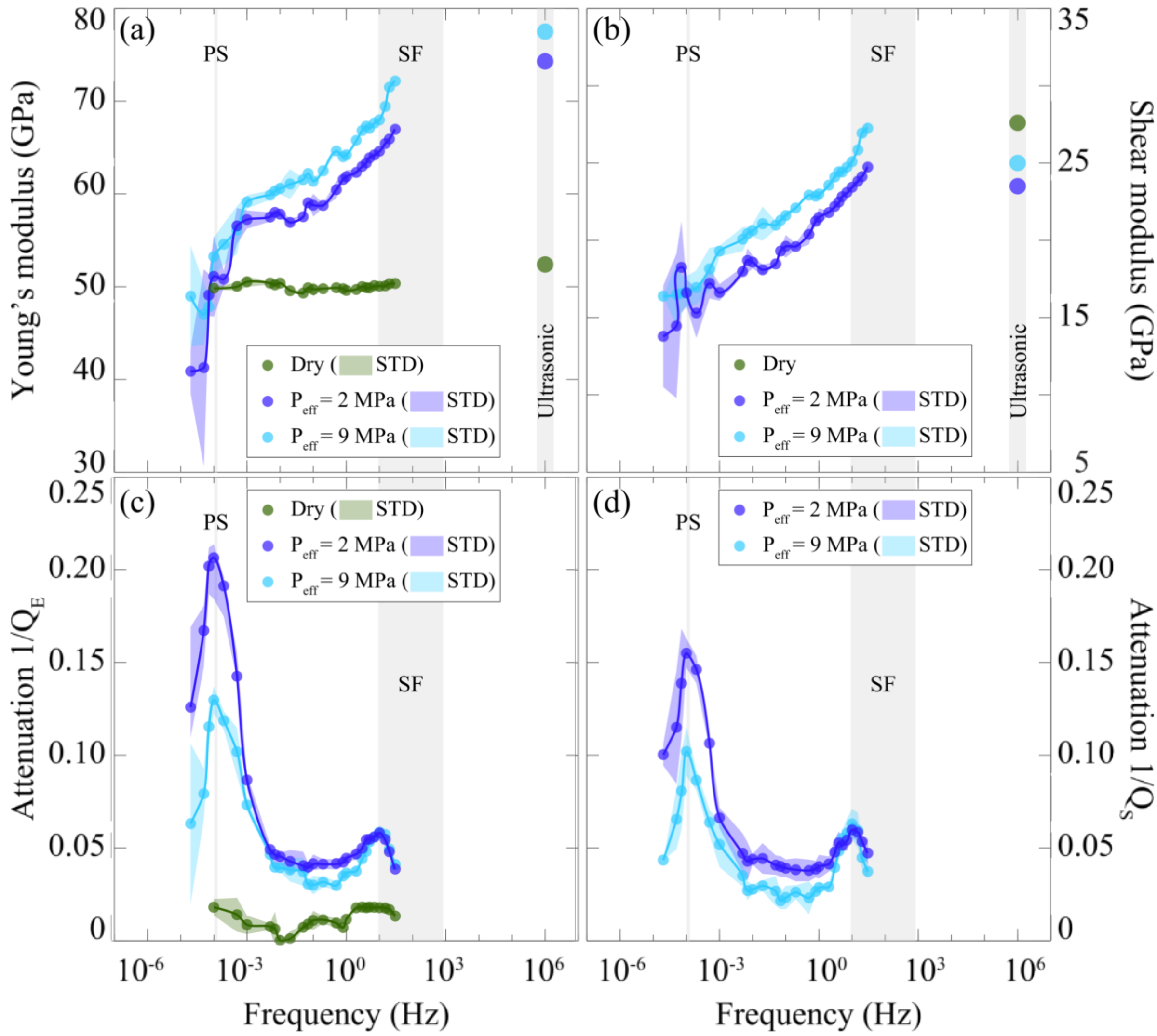


Figure 4 Results for attenuation and elastic moduli as a function of frequency for the Orocopia schist. Green markers indicate dry conditions at 10 MPa effective pressure, dark blue and light blue markers indicate saturated conditions at low (2 MPa) and high (9 MPa) effective pressures, respectively. Shading indicates the standard deviation (STD). The gray bars show the predicted characteristic frequencies for patchy saturation (PS) and squirt flow (SF) mechanisms at experimental conditions, which are described and analyzed in the discussion. (a) Young's modulus, E measured under dry and saturated conditions and compared to published moduli at ultrasonic frequency 1×10^6 Hz from Fliedner and French (2021). (b) Shear modulus G modulus measured under saturated conditions and compared to published moduli at ultrasonic frequency 1×10^6 Hz in Fliedner and French (2021), (c) Attenuation $1/Q_E$ measured under dry and saturated conditions, and (d) Attenuation $1/Q_S$ measured under saturated conditions.

modulus is $\sim 0.012 \pm 0.003$ at all frequencies tested (Figure 4c). In contrast, under saturated conditions, $1/Q_E$ is both higher than under dry conditions and dependent on frequency. For instance, at P_{eff} = 2 MPa, the minimum in $1/Q_E$ is $\sim$ 0.040 ($\pm$ 0.001) between 0.02 and 0.5 Hz, which is about 4 times higher than attenuation under dry conditions. We measure two peaks in $1/Q_E$ under saturated conditions and these are centered at frequencies of 1×10^{-4} Hz and 10 Hz. In addition, the magnitude of $1/Q_E$ decreases with increasing effective pressure. For instance, the peak in attenuation ($1/Q_E$) at 1×10^{-4} Hz is $\sim$ 0.210 at P_{eff} = 2 MPa, whereas it is $\sim$ 0.130 at P_{eff} = 9 MPa. Similarly, the magnitude of $1/Q_S$ is frequency-dependent under saturated conditions and the attenuation is centered at the same frequencies (1×10^{-4} and 10 Hz) as for $1/Q_E$ (Figure 4d). At the 1×10^{-4} Hz peak, $1/Q_S$ is lower than $1/Q_E$ at the same peak by 0.053. At the higher frequency peak (10 Hz), the magnitudes of $1/Q_S$ and $1/Q_E$ are the same within the resolution of our measurements (1×10^{-3}). We observe a constant attenuation between 0.001 and 0.1 under saturated conditions, which corresponds to the frequency range over which the elastic moduli increase linearly. Shear modulus attenuation also decreases with increasing effective pressure, with $1/Q_S$ at an effective pressure of 2 MPa greater than at 9 MPa by 0.021 on average (Figure 4d).

5 Discussion

5.1 Wave-induced Fluid Flow

The presence of pore water and the magnitude of effective pressure both have a measurable effect on the elastic moduli and attenuation of the Orocopia schist. Relative to dry conditions, the attenuation under saturated conditions appears to be increased by a constant value (~ 0.003) with two superimposed frequency dependent peaks (Figure 4c). This constant increase in baseline attenuation is most clearly seen at frequencies between 2×10^{-2} and 3 Hz. As is the case for processes that result in near-constant attenuation, the elastic moduli also clearly increase near-linearly with frequency over this range (Figure 4) (Liu et al., 1976; Kjartansson, 1979). This background attenuation and linear increase in modulus do not arise from wave-induced fluid flow and are not evaluated in detail here.

Peaks in attenuation are frequently interpreted as being caused by wave-induced fluid flow in the rock. In our previous work, we showed that the patchy saturation and the squirt flow mechanisms are consistent with the positions of the peak in attenuation (Figure 4), although we did not evaluate their magnitude, which requires that we consider the dispersion of the elastic moduli (Fliedner and French, 2023j).

The patchy saturation mechanism describes mesoscopic flow between two non-mixing fluids, such as air and water, coexisting in the pore network and forming heterogeneous saturation (White, 1975; Cleary, 1978; Schmitt et al., 1994; Pride et al., 2004). The characteristic frequency of the attenuation peak for patchy saturation (f_{patchy}) is (Cleary, 1978):

$$f_{patchy} = \frac{4kK_d}{\eta L^2} \tag{6}$$

where k is permeability, K_d is the drained bulk modulus, η is water viscosity, and L^2 is the length scale of saturation heterogeneity In Fliedner and French (2023j) we determined that the rock permeability must be $k \sim 1 \times 10^{-21}$ m^2 to explain the position of the peak at 1×10^{-4} Hz for a water viscosity of 10^{-3} $Pa \cdot s$, a drained bulk modulus of 61 GPa which we measured at 2.5 MPa effective pressure under dry conditions, and an assumed length scale of heterogeneous saturation equal to the sample length, $L \sim$ 53 mm. We assumed a constant water viscosity given the negligible temperature variation measured during the experiments. The estimated permeability (1×10^{-21} m^2) is 2 orders of magnitude lower than the permeability measured with a gas pore fluid (4×10^{-19} m^2) and 1 order of magnitude lower than the lower bound of permeability that we estimated by taking into account the Klinkenberg effect (4×10^{-20} m^2). This discrepancy may be attributed to the difference in the experimental conditions at which permeability and attenuation were measured. During the forced oscillations, there is an axial stress of 10 MPa, while the permeability was measured under isotropic stress conditions (Zoback and Byerlee, 1975). Thus, we conclude that the low frequency peak at 1×10^{-4} Hz is generally consistent with the patchy saturation mechanism due to air bubbles trapped in the low permeability pore network (White, 1975; Schmitt et al., 1994; Pride et al., 2004).

The squirt flow mechanism describes fluid flow occurring at the pore scale from thin elongated pores into equant pores (Mavko and Nur, 1975; O'Connell and Budiansky, 1977; Dvorkin et al., 1994). Thin elongated pores are compressible and close with a small magnitude of applied stress whereas equant pores are much less compressible and remain open at the same stresses. Thus, when an elastic wave propagates through the rock, compliant pores close in response and fluid is pushed into the more equant pores (Mavko and Nur, 1975; Dvorkin et al., 1994). Pore geometry is described by the aspect ratio (α), where for an idealized ellipsoidal pore geometry α is the ratio between the short and long axes. The characteristic frequency of the attenuation peak for squirt flow (f_{squirt}) is (O'Connell and Budiansky, 1977):

$$f_{squirt} = \frac{K_s \alpha^3}{\eta} \tag{7}$$

where η is again the viscosity of water and K_s is the bulk modulus of the solid phase, absent pores. We estimated f_{squirt} from the viscosity of water $\eta = 10^{-3}$ $Pa \cdot s$ and our previous measurements on the Orocopia schist. We take the bulk modulus of the solid phase to be K_s = 75 GPa, which we measured under dry conditions and at a confining pressure of 100 MPa, where most porosity is closed (Fliedner and French, 2021, 2023j). We estimated the aspect ratio of the most compliant pore shape using

$$\alpha = \frac{2(1-\nu^2)P_{clos}}{E_0} \tag{8}$$

(Mavko et al., 2020, therein Equation 2.10.71), where ν is Poisson's ratio, P_{clos} is the pressure required to close a pore with aspect ration α and E_0 is the Young's modulus of the solid phase. For the Orocopia schist, we measured a Poisson's ratio of 0.25 and a Young's modulus of 81 GPa at a confining pressure of 100 MPa and under dry conditions. We can assume that the most compliant pores that are open at a given pressure are those just below their closure pressure P_{clos}. As a result, at effective pressures of 2 and 9 MPa, we find the most compliant pore shapes to be $\alpha \sim 4.5 \times 10^{-5}$ and $\alpha \sim 2.1 \times 10^{-4}$, respectively (Fliedner and French, 2023j). If we use these aspect ratios in Equation 7, the resulting characteristic frequencies are 7.5 and 680 Hz for the peaks in attenuation at 2 and 9 MPa effective pressure (Fliedner and French, 2023j). These predicted peaks are close to the position of the high frequency peak (10 Hz), particularly our estimate for 2 MPa effective pressure.

Although we show that the positions of the peaks in attenuation are generally consistent with the patchy saturation and squirt flow mechanisms, these are imprecise assessments of the mechanisms that control attenuation. We use viscoelastic models to assess whether the frequency-dependent attenuation and elastic moduli of the Orocopia schist can be quantitatively explained by the patchy saturation and squirt flow mechanisms.

5.2 Linear Viscoelastic Models

We employ models that assume the system consisting of the rock and pore fluid behaves as a viscoelastic material (O'Connell and Budiansky, 1977). The viscoelasticity of the saturated rock arises from the combined viscous behavior of the fluid and the elastic behavior of the rock frame, assuming negligible viscous deformation of the mineral constituents. Numerically, the complex modulus $M^*(\omega)$ defines the anelastic modulus of a viscoelastic material with elastic (real part M') and viscous (imaginary part, M'') components as:

$$M^*(\omega) = M' + iM'' \tag{9}$$

We use the Maxwell representation of the Standard Linear Solid (SLS) model, also known as a Zener model, to investigate the causality between attenuation and dispersion in the Orocopia schist (Zener and Siegel, 1949). This model assumes a single viscous mechanism with viscosity η_0, that dissipates energy in a short period of time. The mechanical behavior can be visualized using a circuit of springs to represent elastic deformation and dashpots to represent viscous deformation (Figure 5a). The Maxwell model consists of two parallel systems. The first is the Maxwell arm, which operates at the low-frequency limit and is composed of a spring of stiffness M_0 and a dashpot having the viscosity of the fluid, η_0, connected in series. At the high-frequency limit, the undrained elastic modulus, M_∞, controls the mechanical behavior of the rock since the fluid does not have enough time to flow. As a result, fluid viscosity has no effect on the mechanical behavior, and the second arm of the model is composed of a single spring. This model predicts that the attenuation ($1/Q$) has a Debye absorption peak centered at a characteristic relaxation time, τ_c. The characteristic relaxation time is the inverse of the characteristic angular frequency $\omega_c = 2\pi f_c$, with f_c being the characteristic frequency of the viscous mechanisms, in our case for either the patchy saturation (f_{patchy}) or squirt flow (f_{squirt}) mechanisms (Figure 5b). The elastic modulus, M', increases with increasing angular frequency ω, from a low-frequency limit, M_0, to a high-frequency limit, M_∞. We use the formulation of the SLS model presented in (Mavko et al., 2020, therein Table 3.8.1):

$$M^*(\omega) = \frac{M_0 + iM_\infty\left(\dfrac{\omega}{\omega_c}\right)^2}{1+\left(\dfrac{\omega}{\omega_c}\right)^2} \tag{10}$$

From the SLS model, $1/Q$ is determined as the ratio of the imaginary (M'') to real (M') parts of $M^*(\omega)$:

$$1/Q(\omega) = \frac{M''}{M'} = \tan(\phi) \tag{11}$$

The attenuation ($1/Q$) has a maximum magnitude that is proportional to the relaxation strength, Δ, which is the relative difference between the elastic moduli M_∞ and M_0 and occurs at a frequency ω_c:

$$1/Q_{max} = \frac{1}{2}\Delta = \frac{1}{2}\frac{M_\infty - M_0}{\sqrt{M_\infty M_0}} \tag{12}$$

with the relaxation strength given by (Lakes, 2009, Equation 2.50):

$$\Delta = \frac{M_\infty - M_0}{\sqrt{M_\infty M_0}} \tag{13}$$

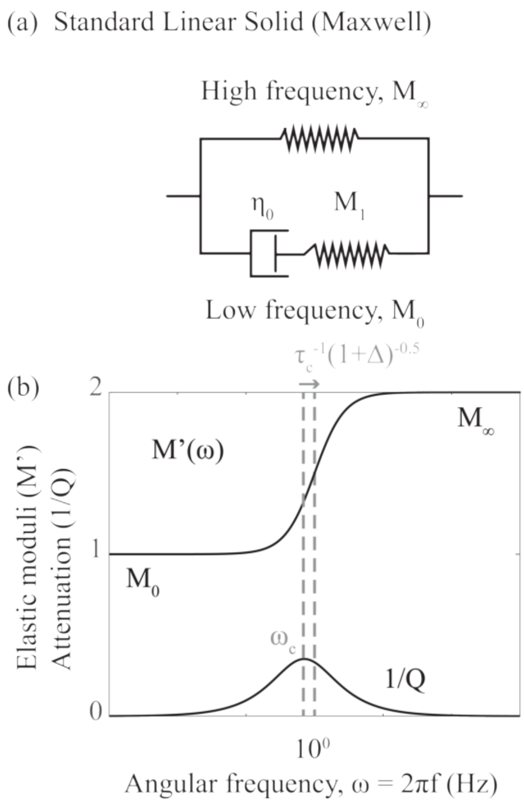


Figure 5 Diagram of the Maxwell representation of a Standard Linear Solid (SLS) model used to describe the viscoelastic behavior of the schist (a) The spring-dashpot analog of the SLS model with two systems in parallel. The first branch is a Maxwell solid that contains a spring (M_1) and a dashpot (η) in series, which represents the elastic and viscous components at the low-frequency limit. The first branch results in the viscoelastic deformation that occurs when there is fluid flow. The second branch is the elastic component at the high-frequency limit (M_∞) when no fluid flow occurs. (b) The SLS model predicts a non-linear increase in elastic moduli from M_0 at the low-frequency limit to M_∞ at the high-frequency limit. The increase in elastic moduli is centered at ω_c which is the characteristic angular frequency of the controlling fluid flow mechanism. The corresponding attenuation ($1/Q$) is a single Debye peak with an amplitude $1/Q_{max}$.

The increase in elastic modulus from the low to high frequency limits is centered near the peak in attenuation at ω_c but is shifted to higher frequencies by $1/(\tau_c\sqrt{(1+\Delta)})$ (Figure 5b).

We estimated M_0 and M_∞ from our experimental measurements of elastic moduli and used the characteristic peak frequencies for patchy saturation and squirt flow (f_{patchy} and f_{squirt}) in Equation 10 to then calculate the attenuation of the Young's and shear moduli as a function of frequency using Equation 11. We evaluated the low and high frequency peaks separately by fitting the data separately over two frequency ranges from 5×10^{-5} to 2×10^{-2} Hz and 3 to 20 Hz corresponding to ranges over which we see peaks in attenuation above some background. The elastic moduli increase linearly between attenuation peaks (2×10^{-2} to 3 Hz), consis-

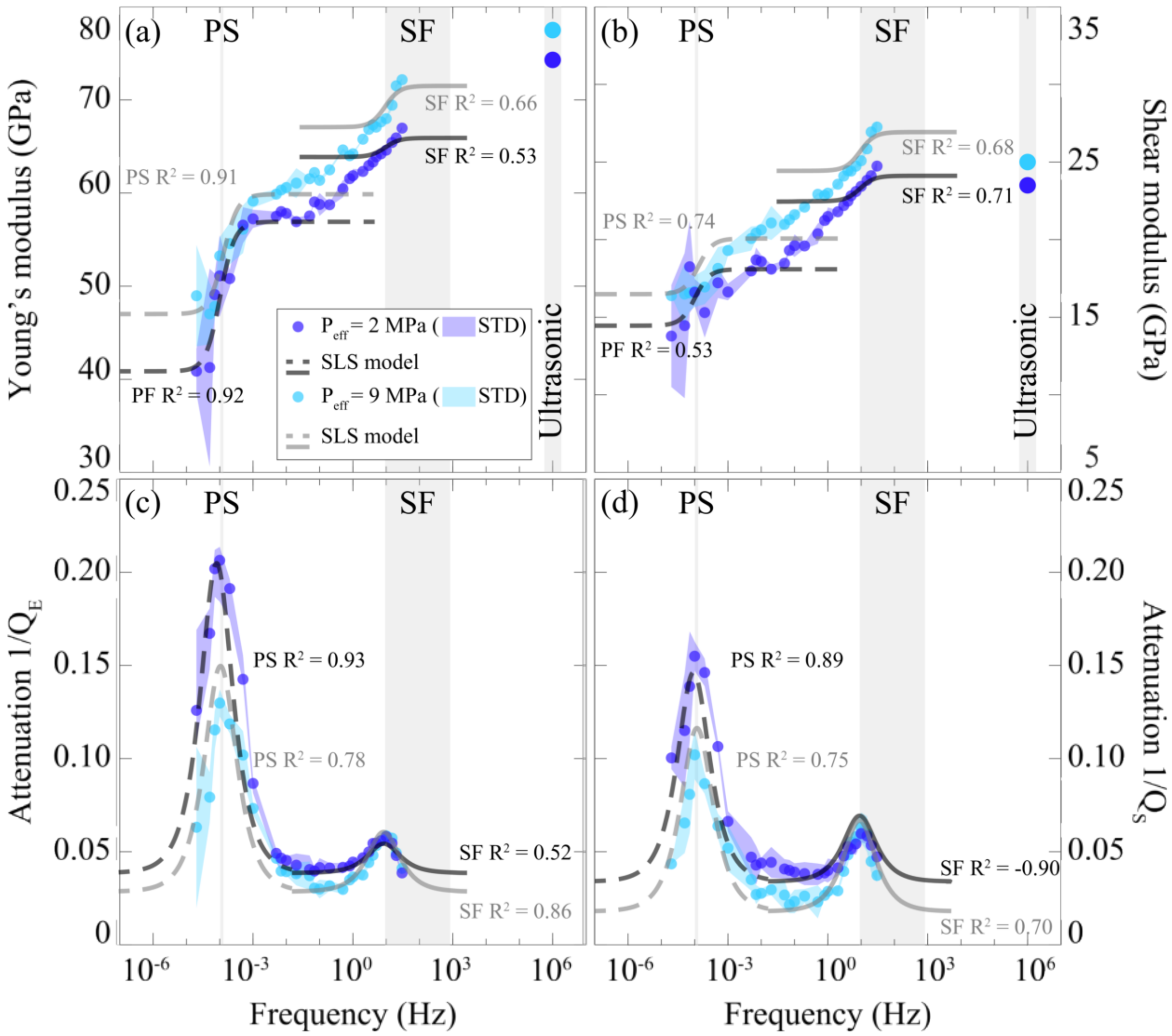


Figure 6 The measurements of frequency-dependent attenuation and elastic moduli under water-saturated conditions as shown in Figure 4 with the results of fitting the Standard Linear Solid (SLS) model to these experimental measurements. Dark blue and light blue colors indicate data at 2 and 9 MPa effective pressures,. The dark and light gray lines indicate model results for the 2 and 9 MPa effective pressures. The model results for patchy saturation are shown with dashed lines and those for squirt flow are shown with solid lines. Gray bars show the predicted characteristic frequencies for the patchy saturation (PS) and squirt flow (SF) mechanisms at experimental conditions. (a) Young's modulus, E with previously reported moduli measured at ultrasonic frequency 1×10^6 Hz from Fliedner and French (2021), (b) shear modulus G with previously reported moduli measured at ultrasonic frequency 1×10^6 Hz from Fliedner and French (2021), (c) Attenuation $1/Q_E$, and (d) Attenuation $1/Q_S$. The goodness of fit of the SLS models to the experimental measurements of attenuation and elasticity were determined with a least square method and are shown, where a perfect fit occurs when R^2 equals 1.

tent with the near-constant attenuation at the same frequencies (Figure 6). Because the changes in elastic moduli due to wave-induced fluid flow are expected to be superimposed on this linear increase, we take M_0 and M_∞ to be the moduli at the low-frequency and high-frequency limits of each peak (Figure 6).

We find that the SLS model generally fits the attenuation data when parameterized with our measurements of elastic moduli and estimate of the peak positions for patchy saturation and squirt flow (Figure 6). Once we calculate attenuation as a function of frequency from Equation 11, we evaluate the goodness of fit between the model prediction and our data (Figure 6). Overall, the goodness of fit is best for the Young's modulus, lower effective pressure (higher fluid pressure), and the low-frequency peak corresponding to patchy saturation. For instance, for the low-frequency peak and Young's modulus, R^2 was greater than 0.78 for both effective pressures (Figure 6a and c). The model for patchy saturation also fits our measurements of shear modulus with an R^2 of 0.53 and attenuation with an R^2 of 0.75 (Figure 6b and d). The goodness of fit of the SLS model to the experimental data is not as strong for the squirt flow mechanism as it is for the patchy saturation mechanism. At 9 MPa effective pressure, the model is in agreement with the measured Young's modulus and related attenuation ($R^2 > 0.58$), but the goodness of fit is poor for the shear modulus at P_{eff} = 2 MPa according to the coefficient of determination $R^2 < 0$. The lower quality of data resulting from lower resolution LVDTs and smaller strain in the radial direction could explain the poor goodness of fit for the shear modulus attenuation. If we calculate the difference between the measured and model $1/Q_E$ at 10 Hz, we find a value of ~ 0.009, consistent with previous attempts to fit SLS models to experimental data (Pimienta et al., 2016b, 2017; Borgomano et al., 2019; Sun et al., 2020).

Because of uncertainty in the aspect ratio of pores, predicting the attenuation and dispersion induced by squirt flow with the Standard Linear Solid model may cause some of the error in the fit. In previous studies of squirt flow, there are visual differences between model results and experimental data (Pimienta et al., 2016b, 2017; Borgomano et al., 2019; Sun et al., 2020), consistent with our relatively low coefficients of determination for the squirt flow mechanism, particularly at low effective stress (Figure 6). One contribution to error may be that the model considers only pores of a single aspect ratio, rather than a distribution of aspect ratios as may occur in rocks. Having a range of pore aspect ratios can cause a more complex squirt flow attenuation signal in a rock such as a broad peak spanning multiple frequencies (Anderson and Archambeau, 1964; Liu et al., 1976) or multiple distinct peaks at different characteristic frequencies (Anderson and Archambeau, 1964; Borgomano et al., 2019). Because we observed a single peak with a relatively narrow width, we used the squirt flow equation for a single aspect ratio. To directly measure the aspect ratio requires precise measurements of both pore diameter and length. Measurement of pore diameters is straightforward with porosimetry methods, but accurately assessing pore length is difficult due to a lack of reliable experimental techniques. Any errors in diameter or length measurement can significantly impact the predicted characteristic frequency of squirt flow because the aspect ratio is cubed in Equation 7. For example, A 10$\%$ margin of error in aspect ratio, gives α between 5.0×10^{-5} and 4.1×10^{-5} at 2 MPa effective pressure and results in a predicted range of characteristic frequencies of 5 and 9 Hz.

5.3 Controls of lithology on attenuation and dispersion

This paper is the first to quantify the attenuation and dispersion of a schist, to the best of our knowledge (Subramaniyan et al., 2014; Rorheim, 2022). We compare our results to other lithologies for which data are available, such as shales (Mikhaltsevitch et al., 2020; Delle Piane et al., 2014), sandstones (Pimienta et al., 2015, 2016b, 2017, 2021; Borgomano et al., 2020; Tisato and Madonna, 2012; Tisato et al., 2015; Madonna and Tisato, 2013), and limestones (Borgomano et al., 2017, 2019). To understand how viscoelasticity is controlled by patchy saturation and squirt flow mechanisms, we discuss the similarities between the rocks, including the relative permeability and pore geometry (Pimienta et al., 2021; Brace, 1977).

Reports of energy dissipation due to patchy saturation are most common in rocks with low permeability and thin elongated pores despite large differences in reported elastic moduli (Pimienta et al., 2021; Cleary, 1978; White, 1975). We find the characteristic frequency f_{patchy} for the Orocopia schist ($\sim 1 \times 10^{-4}$ Hz) to be similar to the predicted characteristic frequency for the Goldwyer shale (2×10^{-4} Hz) (Delle Piane et al., 2014). The Orocopia schist and Goldwyer shale have a similar estimated water permeability ($\sim 10^{-21}$ m^2) and sheet silicate content (>50 $\%$) (Delle Piane et al., 2014; Fliedner and French, 2023j). In some rocks, patchy saturation is also estimated to be important at higher frequencies than we observe, with a f_{patchy} of 40 Hz for a different shale (Mikhaltsevitch et al., 2020) and$\sim$ 300 Hz for a limestone (Borgomano et al., 2017), which are also lithologies with abundant elongated pores. Since both the Orocopia schist and Goldwyer shale have low permeability and small pore throats, bubbles of air can become trapped in their pore network (Pride et al., 2004). When a pressure wave is propagating through, the non-mixing air bubbles flow through the pores due to capillary forces (Schmitt et al., 1994; Delle Piane et al., 2014; Tisato et al., 2015). Thus, the occurrence of non-mixing patches of fluids may be more common in rocks rich in sheet silicates and thin elongated pores than other lithologies (Cleary, 1978; O'Connell and Budiansky, 1977). However, the magnitudes of the attenuation peak in the Orocopia schist and limestone ($1/Q_E >$ 0.13 in Figure 4) are more than double the attenuation peak in the Goldwyer shale ($1/Q_E <$0.07) (Borgomano et al., 2017; Delle Piane et al., 2014; Mikhaltsevitch et al., 2020). This might be caused by differences in the relaxation strength (Equation 13) which links the difference in the Young's modulus and energy dissipation and is consistent with the nearly one order of magnitude difference in elastic moduli between the Goldwyer shale (<7 GPa) and the Orocopia schist (50 GPa) (Figure 6a). Comparisons between experimental data and models that predict attenuation magnitude due to patchy saturation are not common. However, the agreement that we see between experimental data and the model (R^2 > 0.78 in Figure 6a and c) has also been observed in limestone (Borgomano et al., 2017).

The squirt flow mechanism is also a relatively common mechanism of attenuation, particularly in rocks with thin elongated pores. Squirt flow has been interpreted as occurring in several sandstones (Tisato and Madonna, 2012; Madonna and Tisato, 2013; Pimienta et al., 2015, 2017; Subramaniyan et al., 2015; Sun et al., 2020), as well as a thermally cracked limestone (Borgomano et al., 2019), and the frequency of the attenuation peak is strongly dependent on pore shape. The attenuation peak for squirt flow occurs at 10 Hz in the Orocopia schist and most sandstones display this peak between 0.1 and 300 Hz (Pimienta et al., 2017; Borgomano et al., 2019; Subramaniyan et al., 2015; Sun et al., 2020), although it can occur above 1000 Hz due to the typically higher aspect ratio pores in sandstones (Pimienta et al., 2016b; Tisato and Madonna, 2012; Yin et al., 2019). Because of the strong dependence on aspect ratio (Equation 7), the squirt flow mechanism is typically identified when α is $\sim 10^{-4}$ or smaller (Pimienta et al., 2017; Borgomano et al., 2019; Subramaniyan et al., 2015; Sun et al., 2020; Fliedner and French, 2023j). To emphasize the necessity of thin elongated pores even further, Borgomano et al. (2019) only observed evidence of the squirt flow mechanism in limestone after it had been thermally cracked. Despite the fact that both modeling and experimental studies underscore the necessity of elongated pores for squirt flow (Adelinet et al., 2011; Pimienta et al., 2016b), the lack of measurements on schist meant that it was previously unknown whether squirt flow occurs in these rocks at frequencies below 100 Hz (Delle Piane et al., 2014; Mikhaltsevitch et al., 2020).

Rocks rich in phyllosilicates have unique properties that enhance attenuation and dispersion through their influence on microstructure. For instance, rocks rich in phyllosilicates like the Orocopia schist often have a strong shape preferred orientation of grains that then leads to thin and elongated pores aligned with the foliation (Fliedner and French, 2021). As a result, both the permeability and aspect ratios are impacted by the presence of phyllosilicates resulting in high compressibility normal to foliation and high attenuation and dispersion. In addition, phyllosilicate-rich rocks often have weaker fracture strength along their cleavage planes and parallel to foliation, which can cause additional microfracture and macrofracture porosity (Escartín et al., 1997). These macroscopic fractures have a similar shape to thin elongated pores and may cause additional attenuation and dispersion. Specifically, higher microfracture densities adjacent to faults or the presence of larger fractures that are not sampled at the laboratory scale may increase the magnitude of attenuation or cause an additional attenuation peak. Thus, our measurements can be considered a lower bound on the natural system.

5.4 Dispersive seismic velocities and attenuation in subduction zones

Because seismic tomography is the primary tool used to infer potential pore fluid overpressure along the subduction plate boundary, we use our results to evaluate tomographic evidence for fluid pressurization at depths of 25-40 km. In several subduction zones, these depths coincide with regions of low wave velocities, with P-wave velocities (V_P) of 5.0–6.5 km/s and S-wave velocities (V_S) of 2.0–3.2 km/s, and of high V_P/V_S of 2.0–2.8 (Audet et al., 2009; Calvert et al., 2011, 2020; Delph et al., 2018). To compare our results with tomography, we calculate the wave velocities from our measurements of the elastic moduli (Figure 4). The P-wave velocity is determined as $V_P \sim \sqrt{C_{33}/\rho}$ and the S-wave velocity as $V_S \sim \sqrt{G/\rho}$ where C_{33} is the stiffness normal to foliation, G the shear modulus (Figure 4), and ρ the density of the Orocopia schist determined in Fliedner and French (2021). For the Orocopia schist, the Young's modulus normal to the foliation is 30 % lower than C_{33} at an effective pressure of $\sim$ 2 MPa (Fliedner and French, 2021). At a frequency of 1 Hz and an effective pressure of 2 MPa, the Young's modulus is 62 GPa resulting in C_{33} of 81 GPa and the shear modulus is 22 GPa (Figure 4. The resulting estimates of P-wave and S-wave velocities are $V_P \sim 5.2$ km/s and $V_S \sim 2.7$ km/s and their ratio is $V_P/V_S \sim 1.9$. Thus, our approximated velocities at low effective pressure conditions are consistent with the low velocities and high V_P/V_S observed in seismic tomography at 20-45 km in subduction zones (Audet et al., 2009; Calvert et al., 2011, 2020; Delph et al., 2018). We infer that pore fluid overpressurization in metapelites adjacent to the megathrust is consistent with the interpretations of seismic tomography and that the in-situ velocities are most likely dispersive.

Application of our attenuation results requires consideration of differences in laboratory and geologic conditions. Under water saturated conditions, measurements of attenuation are significant, with $1/Q_E$ consistently greater than 0.03. As evaluated in (Fliedner and French, 2023j), the primary consideration when extrapolating to geologic conditions is the effect of fluid viscosity on the positions of the attenuation peaks (Equations 6 and 7). As depth increases, water viscosity decreases from 1×10^{-3} $Pa \cdot s$ at the near-surface to 1×10^{-4} $Pa \cdot s$ at 500 °C (Audetat and Keppler, 2004). When peak position is adjusted to account for the supercritical nature of water at depths of 25-40 km, this increases the characteristic frequencies of wave-induced flow mechanisms (Fliedner and French, 2023j). The adjusted peaks in $1/Q_E$ and $1/Q_S$ for squirt flow coincide with the range of depleted frequencies seen in low-frequency earthquakes (2-8 Hz to 30 Hz) (Farge et al., 2020; Supino et al., 2020) and part of the range of characteristic frequencies observed in very low-frequency earthquakes (0.1 Hz - 30 Hz) (Obana and Kodaira, 2009). In addition, the magnitude of attenuation at these frequencies is sufficient to cause the depletion of these high-frequency waves. While we predict that squirt flow impacts seismic waves significantly, patchy saturation is unlikely unless the permeability is higher at geologic conditions. We estimate the permeability that would be required for the characteristic frequency of patchy saturation to fall within the seismic frequency range (0.1 to 30 Hz). If we assume a bulk modulus similar to Orocopia schist and supercritical water viscosity at in-situ conditions, Equation 6 indicates that a permeability of 1×10^{-19} to 1×10^{-17} m^2 would be necessary to shift the peak to seismic frequencies. These values are high but not implausible for metamorphic rocks, and potentially could result from microfracture damage (Johnson, 1983; Katayama et al., 2012). As a result, models show that our laboratory measurements of attenuation due to squirt flow can explain the limited frequency range of earthquakes (Fliedner and French, 2023j).

We evaluated the role of fluid viscosity on attenuation at geologic conditions, but lithology is likely an important control on the importance of attenuation and dispersion as well. For instance, in warm subduction zones the megathrust typically passes through greenschist facies, whereas cold subduction zones pass through blueschist and eclogite facies (Condit et al., 2020). Thus, although metapelites will undergo some metamorphism, they are expected to be phyllosilicate rich even along colder subduction paths. As we documented for the Orocopia schist, pore structure which is largely controlled by aligned phyllosilicates is the strongest control on bulk rock compliance and attenuation. Thus, although constraining the importance of lithology on attenuation and dispersion would require additional laboratory measurements and microstructural observations of blueschist facies metapelites, we do not predict significant differences in the viscoelastic properties at seismic frequencies.

6 Conclusions

We used the forced oscillation technique to measure the dispersion of elastic moduli in the Orocopia schist at seismic frequencies (2×10^{-5} - 30 Hz) and then used our previous measurements of attenuation and Standard Linear Solid models to evaluate the mechanisms that control attenuation and dispersion. We find that the Young's and shear moduli are dispersive and the magnitudes depend on both water saturation and effective pressure. The Standard Linear Solid models fit the experimental data well and demonstrate that the dispersion is exclusively related to two peaks in attenuation controlled by wave-induced fluid flow. In addition, we show that the patchy saturation and the squirt flow mechanisms of attenuation and dispersion can be described using a single permeability and a single pore geometry, despite the complex microstructure of a schist. In both the schist and other lithologies, the predominance of thin elongated pores is the primary microstructural characteristic controlling the magnitude of dispersion. As a result, attenuation and dispersion are significant and can be modeled for geologic conditions in subduction zones where lithologies rich in phyllosilicates are present.

Acknowledgements

We would like to express our gratitude to our three reviewers: Dave Healy, Nir Badt, and an anonymous reviewer, for their time and effort in reviewing the manuscript. Their valuable comments and suggestions have greatly improved the quality of the manuscript. This research project has been funded by ACS-PRF grant 59440-DN18 to MF and by the David Worthington Named Grant from the AAPG Foundation Grants-In-Aid program to CF.

Data and code availability

The data are available on Zenodo at the following links:

- https://doi.org/10.5281/zenodo.7545104
- https://doi.org/10.5281/zenodo.7545133
- https://doi.org/10.5281/zenodo.7545137
- https://doi.org/10.5281/zenodo.7545181
- https://doi.org/10.5281/zenodo.7528910
- https://doi.org/10.5281/zenodo.7529047
- https://doi.org/10.5281/zenodo.7529147
- https://doi.org/10.5281/zenodo.7529120
- https://doi.org/10.5281/zenodo.7529179

Competing interests

The authors declare no conflict of interest.

References

Adelinet, M., Fortin, J., Guéguen, Y., Schubnel, A., and Geoffroy, L. Frequency and fluid effects on elastic properties of basalt: Experimental investigations. *Geophysical Research Letters*, 37(2): n/a–n/a, 2010. doi: 10.1029/2009gl041660.

Adelinet, M., Fortin, J., and Guéguen, Y. Dispersion of elastic moduli in a porous-cracked rock: Theoretical predictions for squirt-flow. *Tectonophysics*, 503(1-2):173–181, 2011. doi: 10.1016/j.tecto.2010.10.012.

Anderson, D. L. and Archambeau, C. The anelasticity of the earth. *Journal of Geophysical Research*, 69(10):2071–2084, 1964.

Angiboust, S., Kirsch, J., Oncken, O., Glodny, J., Monié, P., and Rybacki, E. Probing the transition between seismically coupled and decoupled segments along an ancient subduction interface. *Geochemistry, Geophysics, Geosystems*, 16(6):1905–1922, 2015. doi: 10.1002/2015gc005776.

Audet, P., Bostock, M. G., Christensen, N. I., and Peacock, S. M. Seismic evidence for overpressured subducted oceanic crust and megathrust fault sealing. *Nature*, 457(7225):76–8, 2009. doi: 10.1038/nature07650.

Audetat, A. and Keppler, H. Viscosity of fluids in subduction zones. *Science*, 303(5657):513–6, 2004. doi: 10.1126/science.1092282.

Batzle, M. L., Han, D.-H., and Hofmann, R. Fluid mobility and frequency-dependent seismic velocity — Direct measurements. *Geophysics*, 71(1):N1–N9, 2006. doi: 10.1190/1.2159053.

Bernabé, Y. and Revil, A. Pore-scale heterogeneity, energy dissipation and the transport properties of rocks. *Geophysical Research Letters*, 22(12):1529–1532, 1995. doi: https://doi.org/10.1029/95GL01418.

Biot, M. A. Theory of Propagation of Elastic Waves in a Fluid-Saturated Porous Solid. II. Higher Frequency Range. *The Journal of the Acoustical Society of America*, 28(2):179–191, 1956. doi: 10.1121/1.1908241.

Borgomano, J. V. M., Pimienta, L., Fortin, J., and Guéguen, Y. Dispersion and attenuation measurements of the elastic moduli of a dual-porosity limestone. *Journal of Geophysical Research: Solid Earth*, 122(4):2690–2711, 2017. doi: 10.1002/2016jb013816.

Borgomano, J. V. M., Pimienta, L. X., Fortin, J., and Guéguen, Y. Seismic Dispersion and Attenuation in Fluid-Saturated Carbonate Rocks: Effect of Microstructure and Pressure. *Journal of Geophysical Research: Solid Earth*, 124(12):12498–12522, 2019. doi: 10.1029/2019jb018434.

Borgomano, J. V. M., Gallagher, A., Sun, C., and Fortin, J. An apparatus to measure elastic dispersion and attenuation using hydrostatic- and axial-stress oscillations under undrained conditions. *Rev Sci Instrum*, 91(3):034502, 2020. doi: 10.1063/1.5136329.

Bostock, M. G., Royer, A. A., Hearn, E. H., and Peacock, S. M. Low frequency earthquakes below southern Vancouver Island. *Geochemistry, Geophysics, Geosystems*, 13(11), 2012. doi: 10.1029/2012gc004391.

Bostock, M. G., Thomas, A. M., Savard, G., Chuang, L., and Rubin, A. M. Magnitudes and moment-duration scaling of low-frequency earthquakes beneath southern Vancouver Island. *Journal of Geophysical Research: Solid Earth*, 120(9):6329–6350, 2015. doi: 10.1002/2015jb012195.

Bostock, M. G., Thomas, A. M., Rubin, A. M., and Christensen, N. I.

On corner frequencies, attenuation, and low-frequency earthquakes. *Journal of Geophysical Research: Solid Earth*, 122(1):543–557, 2017. doi: 10.1002/2016jb013405.

Brace, W. F. Permeability from resistivity and pore shape. *Journal of Geophysical Research*, 82(23):3343–3349, 1977. doi: 10.1029/JB082i023p03343.

Brennan, B. J. and Stacey, F. D. Frequency dependence of elasticity of rock—test of seismic velocity dispersion. *Nature*, 268(5617):220–222, 1977. doi: 10.1038/268220a0.

Brown, R. J. S. and Korringa, J. On the Dependence of the Elastic Properties of a Porous Rock on the Compressibility of the Pore Fluid. *Geophysics*, 40(4):608–616, 1975. doi: 10.1190/1.1440551.

Calvert, A. J., Preston, L. A., and Farahbod, A. M. Sedimentary underplating at the Cascadia mantle-wedge corner revealed by seismic imaging. *Nature Geoscience*, 4(8):545–548, 2011. doi: 10.1038/ngeo1195.

Calvert, A. J., Bostock, M. G., Savard, G., and Unsworth, M. J. Cascadia low frequency earthquakes at the base of an overpressured subduction shear zone. *Nat Commun*, 11(1):3874, 2020. doi: 10.1038/s41467-020-17609-3.

Chapman, S., Tisato, N., Quintal, B., and Holliger, K. Seismic attenuation in partially saturated Berea sandstone submitted to a range of confining pressures. *Journal of Geophysical Research: Solid Earth*, 121(3):1664–1676, 2016. doi: 10.1002/2015jb012575.

Cleary, M. P. Elastic and dynamic response regimes of fluid-impregnated solids with diverse microstructures. *International Journal of Solids and Structures*, 14(10):795–819, 1978. doi: 10.1016/0020-7683(78)90072-0.

Condit, C. B. and French, M. E. Geologic Evidence of Lithostatic Pore Fluid Pressures at the Base of the Subduction Seismogenic Zone. *Geophysical Research Letters*, 49(12), 2022. doi: 10.1029/2022gl098862.

Condit, C. B., Guevara, V. E., Delph, J. R., and French, M. E. Slab dehydration in warm subduction zones at depths of episodic slip and tremor. *Earth and Planetary Science Letters*, 552, 2020. doi: 10.1016/j.epsl.2020.116601.

Condit, C. B., French, M. E., Hayles, J. A., Yeung, L. Y., Chin, E. J., and Lee, C. A. Rheology of Metasedimentary Rocks at the Base of the Subduction Seismogenic Zone. *Geochemistry, Geophysics, Geosystems*, 23(2), 2022. doi: 10.1029/2021gc010194.

Delle Piane, C., Sarout, J., Madonna, C., Saenger, E. H., Dewhurst, D. N., and Raven, M. Frequency-dependent seismic attenuation in shales: experimental results and theoretical analysis. *Geophysical Journal International*, 198(1):504–515, 2014. doi: 10.1093/gji/ggu148.

Delph, J. R., Levander, A., and Niu, F. Fluid Controls on the Heterogeneous Seismic Characteristics of the Cascadia Margin. *Geophysical Research Letters*, 45(20), 2018. doi: 10.1029/2018gl079518.

Duffy, W. Acoustic quality factor of aluminium and selected aluminium alloys from 50 mK to 300 K. *Cryogenics*, 42(3-4):245–251, 2002. doi: 10.1016/s0011-2275(02)00021-8.

Dunn, K. Acoustic attenuation in fluid-saturated porous cylinders at low frequencies. *The Journal of the Acoustical Society of America*, 79(6):1709–1721, 1986. doi: 10.1121/1.393232.

Dvorkin, J., Nolen-Hoeksema, R., and Nur, A. The squirt-flow mechanism: Macroscopic description. *Geophysics*, 59(3):428–438, 1994. doi: 10.1190/1.1443605.

Escartín, J., Hirth, G., and Evans, B. Nondilatant brittle deformation of serpentinites: Implications for Mohr-Coulomb theory and the strength of faults. *Journal of Geophysical Research: Solid Earth*, 102(B2):2897–2913, 1997. doi: 10.1029/96jb02792.

Farge, G., Shapiro, N. M., and Frank, W. B. Moment-Duration Scaling of Low-Frequency Earthquakes in Guerrero, Mexico. *Journal of Geophysical Research: Solid Earth*, 125(8), 2020. doi: 10.1029/2019jb019099.

Fliedner, C. and French, M. Forced Oscillations On PMMA (1). https://doi.org/10.5281/zenodo.7545104, 2023a.

Fliedner, C. and French, M. Forced Oscillations On The Orocopia Schist Under Dry Conditions. https://doi.org/10.5281/zenodo.7528910, 2023b.

Fliedner, C. and French, M. Forced Oscillations On Aluminum (1). https://doi.org/10.5281/zenodo.7545133, 2023c.

Fliedner, C. and French, M. Forced Oscillations On The Orocopia Schist Under Saturated Conditions (Effective Stress 9 MPa) (1). https://doi.org/10.5281/zenodo.7529047, 2023d.

Fliedner, C. and French, M. Forced Oscillations On The Orocopia Schist Under Saturated Conditions (Effective Stress 2 MPa) (1). https://doi.org/10.5281/zenodo.7529120, 2023e.

Fliedner, C. and French, M. Forced Oscillations On PMMA (2). https://doi.org/10.5281/zenodo.7545137, 2023f.

Fliedner, C. and French, M. Forced Oscillations On Aluminum (2). https://doi.org/10.5281/zenodo.7545181, 2023g.

Fliedner, C. and French, M. Forced Oscillations On The Orocopia Schist Under Saturated Conditions (Effective Stress 9 MPa) (2). https://doi.org/10.5281/zenodo.7529147, 2023h.

Fliedner, C. and French, M. Forced Oscillations On The Orocopia Schist Under Saturated Conditions (Effective Stress 2 MPa) (2). https://doi.org/10.5281/zenodo.7529179, 2023i.

Fliedner, C. and French, M. E. Pore and Mineral Fabrics Control the Elastic Wave Velocities of Metapelite With Implications for Subduction Zone Tomography. *Journal of Geophysical Research: Solid Earth*, 126(10), 2021. doi: 10.1029/2021jb022361.

Fliedner, C. and French, M. E. Measurements of wave-induced attenuation in saturated metapelite and the band-limitation of low-frequency earthquakes. *AGU Advances*, 4(2), 2023j. doi: 10.1029/2022AV000837.

Gassmann, F. Über die Elastizität poröser Medien. Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich. *GEOPHYSICS*, 96:1–23, 1951.

Gomberg, J., Creager, K., Sweet, J., Vidale, J., Ghosh, A., and Hotovec, A. Earthquake spectra and near-source attenuation in the Cascadia subduction zone. *Journal of Geophysical Research: Solid Earth*, 117(B5):n/a–n/a, 2012. doi: 10.1029/2011jb009055.

Gomberg, J., Agnew, D. C., and Schwartz, S. Y. Alternative source models of very low frequency events. *Journal of Geophysical Research: Solid Earth*, 121(9):6722–6740, 2016. doi: 10.1002/2016jb013001.

Ide, S., Beroza, G. C., Shelly, D. R., and Uchide, T. A scaling law for slow earthquakes. *Nature*, 447(7140):76–79, 2007. doi: 10.1038/nature05780.

Ito, Y., Obara, K., Shiomi, K., Sekine, S., and Hirose, H. Slow earthquakes coincident with episodic tremors and slow slip events. *Science*, 315(5811):503–6, 2007. doi: 10.1126/science.1134454.

Jackson, I. and Paterson, M. S. Shear modulus and internal friction of calcite rocks at seismic frequencies: pressure, frequency and grain size dependence. *Physics of the Earth and Planetary Interiors*, 45(4):349–367, 1987. doi: 10.1016/0031-9201(87)90042-2.

Jacobson, C. E. and Dawson, M. R. Structural and metamorphic evolution of the Orocopia Schist and related rocks, southern California: Evidence for late movement on the Orocopia fault. *Tectonics*, 14(4):933–944, 1995. doi: 10.1029/95tc01446.

Jacobson, C. E., Grove, M., Vućić, A., Pedrick, J. N., and Ebert, K. A. *Exhumation of the Orocopia Schist and associated rocks of south-*

eastern California: Relative roles of erosion, synsubduction tectonic denudation, and middle Cenozoic extension, volume 419, book section 1, pages 1–37. Geological Society of America, 2007. doi: 10.1130/2007.2419(01).

Johnson, G. R. Rock property measurements and analysis of selected igneous, sedimentary, and metamorphic rocks from worldwide localities. Report 2331-1258, USGS, 1983.

Karato, S. Importance of anelasticity in the interpretation of seismic tomography. *Geophysical research letters*, 20(15): 1623–1626, 1993.

Katayama, I., Terada, T., Okazaki, K., and Tanikawa, W. Episodic tremor and slow slip potentially linked to permeability contrasts at the Moho. *Nature Geoscience*, 5(10):731–734, 2012. doi: 10.1038/ngeo1559.

Kjartansson, E. Constant Q-wave propagation and attenuation. *Journal of Geophysical Research*, 84(B9), 1979. doi: 10.1029/JB084iB09p04737.

Kranz, R. L. Microcracks in rocks: A review. *Tectonophysics*, 100 (1-3):449–480, 1983. doi: 10.1016/0040-1951(83)90198-1.

Lakes, R. *Viscoelastic Materials*. Cambridge University Press, 2009. doi: 10.1017/CBO9780511626722.

Lee, T., Lakes, R. S., and Lal, A. Resonant ultrasound spectroscopy for measurement of mechanical damping: Comparison with broadband viscoelastic spectroscopy. *Review of Scientific Instruments*, 71(7):2855–2861, 2000. doi: 10.1063/1.1150703.

Littel, G. F., Thomas, A. M., and Baltay, A. S. Using Tectonic Tremor to Constrain Seismic Wave Attenuation in Cascadia. *Geophysical Research Letters*, 45(18):9579–9587, 2018. doi: 10.1029/2018gl079344.

Liu, H. P., Anderson, D. L., and Kanamori, H. Velocity dispersion due to anelasticity; implications for seismology and mantle composition. *Geophysical Journal International*, 47(1):41–58, 1976. doi: 10.1111/j.1365-246X.1976.tb01261.x.

Madonna, C. and Tisato, N. A new Seismic Wave Attenuation Module to experimentally measure low-frequency attenuation in extensional mode. *Geophysical Prospecting*, 61(2):302–314, 2013. doi: 10.1111/1365-2478.12015.

Mavko, G. and Nur, A. Melt squirt in the asthenosphere. *Journal of Geophysical Research*, 80(11):1444–1448, 1975. doi: 10.1029/JB080i011p01444.

Mavko, G., Mukerji, T., and Dvorkin, J. *The Rock Physics Handbook*. Cambridge University Press, 3rd edition edition, 2020. doi: 10.1017/9781108333016.

Mavko, G. M. and Nur, A. The effect of nonelliptical cracks on the compressibility of rocks. *Journal of Geophysical Research*, 83 (B9), 1978. doi: 10.1029/JB083iB09p04459.

Mikhaltsevitch, V., Lebedev, M., Pervukhina, M., and Gurevich, B. Seismic dispersion and attenuation in Mancos shale – laboratory measurements. *Geophysical Prospecting*, 69(3):568–585, 2020. doi: 10.1111/1365-2478.13056.

Muñoz-Montecinos, J., Angiboust, S., and Garcia-Casco, A. Blueschist-facies paleo-earthquakes in a serpentinite channel (Zagros suture, Iran) enlighten seismogenesis in Mariana-type subduction margins. *Earth and Planetary Science Letters*, 573, 2021. doi: 10.1016/j.epsl.2021.117135.

Nakai, J. S., Sheehan, A. F., Abercrombie, R. E., and Eberhart-Phillips, D. Near Trench 3D Seismic Attenuation Offshore Northern Hikurangi Subduction Margin, North Island, New Zealand. *Journal of Geophysical Research: Solid Earth*, 126(3), 2021. doi: 10.1029/2020jb020810.

Nowick, A. S. and Berry, B. S. *Anelastic Relaxation in Crystalline Solids*. Academic Press, 1972. doi: https://doi.org/10.1016/B978-0-12-522650-9.50005-7.

Obana, K. and Kodaira, S. Low-frequency tremors associated with reverse faults in a shallow accretionary prism. *Earth and Planetary Science Letters*, 287(1-2):168–174, 2009. doi: 10.1016/j.epsl.2009.08.005.

Obara, K. Nonvolcanic deep tremor associated with subduction in southwest Japan. *Science*, 296(5573):1679–81, 2002. doi: 10.1126/science.1070378.

Obara, K. and Kato, A. Connecting slow earthquakes to huge earthquakes. *Science*, 353(6296):253–7, 2016. doi: 10.1126/science.aaf1512.

Oberg, E. and McCauley, C. J. *Machinery's Handbook: Toolbox*. Industrial Press, Inc., 31st edition edition, 2020.

O'Connell, R. J. and Budiansky, B. Seismic velocities in dry and saturated cracked solids. *Journal of Geophysical Research*, 79 (35):5412–5426, 1974. doi: 10.1029/JB079i035p05412.

O'Connell, R. J. and Budiansky, B. Viscoelastic properties of fluid-saturated cracked solids. *Journal of Geophysical Research*, 82 (36):5719–5735, 1977. doi: 10.1029/JB082i036p05719.

Peacock, S. M. Serpentinization and infiltration metasomatism in the Trinity peridotite, Klamath province, northern California: implications for subduction zones. *Contributions to Mineralogy and Petrology*, 95(1):55–70, 1987. doi: 10.1007/bf00518030.

Philippot, P. and Selverstone, J. Trace-element-rich brines in eclogitic veins: implications for fluid composition and transport during subduction. *Contributions to Mineralogy and Petrology*, 106(4):417–430, 1991. doi: 10.1007/bf00321985.

Pimienta, L., Fortin, J., and Guéguen, Y. Experimental study of Young's modulus dispersion and attenuation in fully saturated sandstones. *Geophysics*, 80(5):L57–L72, 2015. doi: 10.1190/geo2014-0532.1.

Pimienta, L., Borgomano, J. V. M., Fortin, J., and Guéguen, Y. Modelling the drained/undrained transition: effect of the measuring method and the boundary conditions. *Geophysical Prospecting*, 64(4):1098–1111, 2016a. doi: 10.1111/1365-2478.12390.

Pimienta, L., Fortin, J., and Guéguen, Y. Effect of fluids and frequencies on Poisson's ratio of sandstone samples. *Geophysics*, 81(2):D183–D195, 2016b. doi: 10.1190/GEO2015-0310.1.

Pimienta, L., Borgomano, J. V. M., Fortin, J., and Guéguen, Y. Elastic Dispersion and Attenuation in Fully Saturated Sandstones: Role of Mineral Content, Porosity, and Pressures. *Journal of Geophysical Research: Solid Earth*, 122(12):9950–9965, 2017. doi: 10.1002/2017jb014645.

Pimienta, L., Quintal, B., and Caspari, E. Hydro-mechanical coupling in porous rocks: hidden dependences to the microstructure? *Geophysical Journal International*, 224(2):973–984, 2021. doi: 10.1093/gji/ggaa497.

Platt, J. P., Xia, H., and Schmidt, W. L. Rheology and stress in subduction zones around the aseismic/seismic transition. *Progress in Earth and Planetary Science*, 5(1), 2018. doi: 10.1186/s40645-018-0183-8.

Pride, S. R., Berryman, J. G., and Harris, J. M. Seismic attenuation due to wave-induced flow. *Journal of Geophysical Research: Solid Earth*, 109(B1), 2004. doi: 10.1029/2003jb002639.

Rorheim, S. On Frequency-Dependent Rock Experiments: A Comparative Review. *arXiv*, 2022. doi: https://doi.org/10.48550/arXiv.2208.03795.

Saltiel, S., Selvadurai, P. A., Bonner, B. P., Glaser, S. D., and Ajo-Franklin, J. B. Experimental development of low-frequency shear modulus and attenuation measurements in mated rock fractures: Shear mechanics due to asperity contact area changes with normal stress. *Geophysics*, 82(2):M19–M36, 2017. doi: 10.1190/geo2016-0199.1.

Sammis, C. G. and Bostock, M. G. A Granular Jamming Model for

Low-Frequency Earthquakes. *Journal of Geophysical Research: Solid Earth*, 126(7), 2021. doi: 10.1029/2021jb021963.

Schmitt, L., Forsans, T., and Santarelli, F. J. Shale testing and capillary phenomena. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 31(5):411–427, 1994. doi: 10.1016/0148-9062(94)90145-7.

Shapiro, N. M., Campillo, M., Kaminski, E., Vilotte, J., and Jaupart, C. Low-Frequency Earthquakes and Pore Pressure Transients in Subduction Zones. *Geophysical Research Letters*, 45(20), 2018. doi: 10.1029/2018gl079893.

Shelly, D. R., Beroza, G. C., Ide, S., and Nakamula, S. Low-frequency earthquakes in Shikoku, Japan, and their relationship to episodic tremor and slip. *Nature*, 442(7099):188–91, 2006. doi: 10.1038/nature04931.

Shelly, D. R., Beroza, G. C., and Ide, S. Non-volcanic tremor and low-frequency earthquake swarms. *Nature*, 446(7133):305–7, 2007. doi: 10.1038/nature05666.

Spencer, J. W. Stress relaxations at low frequencies in fluid-saturated rocks: Attenuation and modulus dispersion. *Journal of Geophysical Research*, 86(B3), 1981. doi: 10.1029/JB086iB03p01803.

Subramaniyan, S., Quintal, B., Tisato, N., Saenger, E. H., and Madonna, C. An overview of laboratory apparatuses to measure seismic attenuation in reservoir rocks. *Geophysical Prospecting*, 62(6):1211–1223, 2014. doi: 10.1111/1365-2478.12171.

Subramaniyan, S., Quintal, B., Madonna, C., and Saenger, E. H. Laboratory-based seismic attenuation in Fontainebleau sandstone: Evidence of squirt flow. *Journal of Geophysical Research: Solid Earth*, 120(11):7526–7535, 2015. doi: 10.1002/2015jb012290.

Sun, C., Tang, G., Fortin, J., Borgomano, J. V. M., and Wang, S. Dispersion and Attenuation of Elastic Wave Velocities: Impact of Microstructure Heterogeneity and Local Measurements. *Journal of Geophysical Research: Solid Earth*, 125(12), 2020. doi: 10.1029/2020jb020132.

Supino, M., Poiata, N., Festa, G., Vilotte, J. P., Satriano, C., and Obara, K. Self-similarity of low-frequency earthquakes. *Sci Rep*, 10(1):6523, 2020. doi: 10.1038/s41598-020-63584-6.

Tanikawa, W. and Shimamoto, T. Comparison of Klinkenberg-corrected gas permeability and water permeability in sedimentary rocks. *International Journal of Rock Mechanics and Mining Sciences*, 46(2):229–238, 2009. doi: 10.1016/j.ijrmms.2008.03.004.

Terzaghi, K., Peck, R., and Mesri, G. *Soil Mechanics in Engineering Practice*. Wiley, 1996.

Tewksbury-Christle, C. M., Behr, W. M., and Helper, M. A. Tracking Deep Sediment Underplating in a Fossil Subduction Margin: Implications for Interface Rheology and Mass and Volatile Recycling. *Geochem Geophys Geosyst*, 22(3):e2020GC009463, 2021. doi: 10.1029/2020GC009463.

Thomas, A. M., Beroza, G. C., and Shelly, D. R. Constraints on the source parameters of low-frequency earthquakes on the San Andreas Fault. *Geophysical Research Letters*, 43(4):1464–1471, 2016. doi: 10.1002/2015gl067173.

Tisato, N. and Madonna, C. Attenuation at low seismic frequencies in partially saturated rocks: Measurements and description of a new apparatus. *Journal of Applied Geophysics*, 86:44–53, 2012. doi: 10.1016/j.jappgeo.2012.07.008.

Tisato, N., Quintal, B., Chapman, S., Podladchikov, Y., and Burg, J. Bubbles attenuate elastic waves at seismic frequencies: First experimental evidence. *Geophysical Research Letters*, 42(10):3880–3887, 2015. doi: 10.1002/2015gl063538.

Toksöz, M. N., Johnston, D. H., and Timur, A. Attenuation of seismic waves in dry and saturated rocks: I. Laboratory measurements. *Geophysics*, 44(4):681–690, 1979. doi: 10.1190/1.1440969.

Walsh, J. B. The effect of cracks on the compressibility of rock. *Journal of Geophysical Research*, 70(2):381–389, 1965. doi: 10.1029/JZ070i002p00381.

Wei, X., Xu, J., Liu, Y., and Chen, X. The slow self-arresting nature of low-frequency earthquakes. *Nat Commun*, 12(1):5464, 2021. doi: 10.1038/s41467-021-25823-w.

White, J. E. Computed Seismic Speeds and Attenuation in Rocks with Partial Gas Saturation. *Geophysics*, 40(2):224–232, 1975. doi: 10.1190/1.1440520.

Winkler, K. and Nur, A. Pore fluids and seismic attenuation in rocks. *Geophysical Research Letters*, 6(1):1–4, 1979. doi: 10.1029/GL006i001p00001.

Yin, H., Borgomano, J. V. M., Wang, S., Tiennot, M., Fortin, J., and Guéguen, Y. Fluid Substitution and Shear Weakening in Clay-Bearing Sandstone at Seismic Frequencies. *Journal of Geophysical Research: Solid Earth*, 124(2):1254–1272, 2019. doi: 10.1029/2018jb016241.

Zener, C. M. and Siegel, S. Elasticity and Anelasticity of Metals. *The Journal of Physical and Colloid Chemistry*, 53(9):1468–1468, 1949. doi: 10.1021/j150474a017.

Zoback, M. D. and Byerlee, J. D. The effect of microcrack dilatancy on the permeability of westerly granite. *Journal of Geophysical Research*, 80(5):752–755, 1975. doi: 10.1029/JB080i005p00752.

doi:10.26443/seismica.v3i1.386
SEISMICA

Influence of outer-rise faults on shallow décollement heterogeneity and sediment flux at the Japan trench

E. Schottenfels [1], **C. Regalla** *[1], **Y. Nakamura** [2]

[1]School of Earth and Sustainability, Northern Arizona University, Flagstaff, AZ, USA, [2]Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Kanagawa, Japan

Author contributions: *Conceptualization*: E. Schottenfels, C. Regalla. *Data Curation*: Y. Nakamura. *Formal Analysis*: E. Schottenfels, C. Regalla. *Funding Acquisition*: C. Regalla. *Investigation*: E. Schottenfels. *Methodology*: E. Schottenfels. *Supervision*: C. Regalla. *Validation*: Y. Nakamura. *Visualization*: E. Schottenfels. *Writing – original draft*: E. Schottenfels. *Writing – review & editing*: E. Schottenfels, C. Regalla, Y. Nakamura.

Abstract We investigate the impact of outer-rise normal fault subduction on the structural evolution of the décollement and frontal prism in a portion of the Japan trench that hosted the 2011 Tohoku earthquake. We use seismic reflection data to map the relative occurrence of sediment accretion, sediment subduction, and frontal tectonic erosion in the shallow portion of the subduction zone and correlate these deformation styles to the magnitude of outer-rise fault throw and incoming plate sediment thickness. These data reveal spatial heterogeneity in the modes of deformation over distances of 5-10 km that necessitate correlative heterogeneity in the geometry and composition of the shallow décollement over similar length-scales. We find that sediment accretion predominantly occurs in regions where incoming plate sediment thickness is greater than fault throw. In these areas, the décollement appears to be non-planar and compositionally homogenous. Conversely, frontal tectonic erosion and slope failures are predominantly observed in regions where fault throw is greater than sediment thickness. In these areas, the décollement may be planar but compositionally heterogeneous. Additionally, spatial variations in near trench slip appear to correlate with the dominant deformation modes, suggesting that both sediment thickness and outer-rise fault throw may be important controls on shallow megathrust behavior.

Non-technical summary We investigate how properties of the subducting plate affect the structure of the shallow subduction zone off the coast of northern Japan, and how this may impact the earthquake potential of the region. We use geophysical reflection data to determine how sediment on top of the oceanic crust, and faults that displace the crust, can allow for sediment to be either scraped off or brought down into the subduction zone. We determine that these processes depend on both how thick the sediment is and how much displacement has occurred along faults. In areas where the sediment is thicker than the fault displacement, sediments are off-scraped and the interface between the upper and lower plate is more uniform in composition. In areas where the faults are larger than sediment thickness, the sediment is subducted, and the interface may have a varied composition. We identify variations in sediment and interface properties that have important implications for how the subduction zone evolves over time and for the style of shallow earthquakes. We suggest that the thickness of the sediment and the size of the faults on the subducting plate can help us understand such earthquake behavior.

Production Editor:
Gareth Funning
Handling Editor:
Randolph Williams
Copy & Layout Editor:
Hannah F. Mark

Signed reviewer(s):
Rebecca Bell
Harold Tobin

Received:
February 18, 2023
Accepted:
December 5, 2023
Published:
January 15, 2024

1 Introduction

The subduction of outer-rise normal faults is a nearly ubiquitous process that plays an important role in the mechanics of the shallow portion of the plate boundary fault (décollement) as lower plate surface roughness can modulate the frictional and fluid properties of the subduction interface (Clift and Vannucchi, 2004; Moore et al., 1986; Morgan et al., 2007; Polet and Kanamori, 2000; Saffer and Tobin, 2011; Tanioka et al., 1997). In many sediment-rich margins, this lower plate roughness is covered with a thick sediment sequence that often allows the décollement to "smooth over" the subducting topography (e.g., Contreras-Reyes et al., 2007; Wang and Bilek, 2014). In contrast, in sediment-poor margins, such as northeastern Japan, outer-rise normal faults produce significant offsets of the subducting oceanic crust and the seafloor outboard of the trench (e.g., Masson, 1991), which impart geometric and compositional heterogeneities that disrupt the continuity and evolution of the décollement (e.g., Polet and Kanamori, 2000; Tanioka et al., 1997). Constraints on the stratigraphic position and geometric evolution of the décollement are necessary for understanding the first-order control on the volumes of subducted versus accreted sediments, and the composition and strength of the shallow décollement (Saffer and Tobin, 2011; Moore et al., 2015; Polet and Kanamori, 2000). However, the factors controlling the evolution of the shallow décollement in response to subducting outer-rise faults remain unclear.

We evaluate how incoming plate sediment thickness

*Corresponding author: Christine.Regalla@nau.edu

and relief across outer-rise normal faults influence the evolution of the décollement and sediment fluxes at the shallow portion of the Japan trench. The Japan trench is a well-imaged margin that has become a type location for studies on outer-rise normal fault subduction and its impacts on sediment fluxes and forearc evolution (e.g., von Huene and Lallemand, 1990; Tanioka et al., 1997; von Huene and Culotta, 1989). Two different scenarios have been proposed for the evolution of the shallow portion of the subduction interface in northeastern Japan. In the first scenario, the development of a planar décollement across subducting horsts and grabens entraps both lower and upper plate sediments within subducted grabens and prevents voluminous sediment accretion (e.g., Hilde, 1983; von Huene et al., 1982; von Huene and Culotta, 1989). This model is supported by seismic reflection data that show sediment slumping into subducting grabens at the trench and a décollement that appears to project across lower plate grabens beneath the frontal prism (e.g., Hilde, 1983; Kodaira et al., 2012; Strasser et al., 2013). In the second scenario, the shallow décollement geometry mimics the subducted seafloor roughness, such that the majority of the incoming sediment section is off-scraped and accreted (Nakamura et al., 2013; Regalla et al., 2019). This model is supported by high-resolution seismic reflection data (Nakamura et al., 2013, 2020) and drill cores collected across the plate boundary interface at the Japan trench (Chester et al., 2013) that show thrust imbrication of sediments near the trench and image an undulating décollement that appears to step over subducting horsts and grabens. Recently collected high-resolution seismic reflection data along the Japan trench show that both modes can occur along the margin (Nakamura et al., 2020). However, it remains unclear which processes control the evolution of the décollement in response to subducting horsts and grabens, the spatial variations in these processes along strike and down dip, and the relationship to the styles of slip along the shallow subduction interface.

Here, we analyze 44 seismic reflection profiles in a 40 x 180 km portion of the margin between 38°N and 40°N, in order to evaluate the relationship between sediment thickness, outer-rise fault throw, and modes of frontal prism deformation in a sediment-starved region that is known to be capable of hosting tsunamigenic slip (Figure 1). We calculate incoming plate sediment thickness and throw across normal faults on the incoming plate and beneath the prism, and identify modes of frontal prism deformation including sediment accretion, partial sediment accretion, complete sediment subduction, and frontal tectonic erosion. Our results show that the relative magnitudes of fault throw and sediment thickness together control modes of prism deformation, sediment flux, and shallow décollement heterogeneity. Sediment accretion is generally observed where sediment thickness is large relative to outer-rise fault throw, whereas frontal tectonic erosion occurs where fault throw is large relative to sediment thickness. These two styles of deformation have different endmember implications for the compositional and geometric heterogeneity of the shallow décollement that appear to correlate with different styles of shallow subduction interface slip. Therefore, sediment thickness and outer-rise fault throw may be important factors for predicting modes of frontal prism deformation, heterogeneity in the shallow subduction interface, and the conditions that may promote or inhibit shallow plate boundary slip.

2 Background

This study is performed along a portion of the NE Japan trench margin where high-density seismic reflection data allow for detailed evaluations of the frontal prism and incoming plate structure. At this location, the Pacific plate subducts beneath northern Honshu at a rate of ~80 mm/yr at the Japan trench (Seno et al., 1993) (Figure 1). The forearc consists of a high-velocity wedge (4-6 km/s P-wave) made of Cretaceous and younger accreted sediments (Tsuru et al., 2002; Kodaira et al., 2017), and a low velocity (2-3.5 km/s P-wave), seismically chaotic frontal prism, located at the seaward edge of the forearc. The prism is a wedge-shaped sedimentary package with a width of ~15–30 km that spans the margin parallel to the trench axis (Tsuru et al., 2002; Kodaira et al., 2017). The wedge appears to be composed of accreted incoming plate sediments derived from biogenic muds off-scraped from the incoming Pacific plate (Nakamura et al., 2013; Regalla et al., 2019), modified by large slope failures (Nakamura et al., 2020).

The incoming Pacific plate has a relatively thin sediment cover (~50–600 m) (Boston et al., 2014; Nakamura et al., 2023) due to limited biologic productivity in the north Pacific gyre and the relatively small volume of trench fill derived from the nearby continent and trench slopes (Moore et al., 2015; Ikehara et al., 2017). The incoming plate stratigraphy recovered at Deep Sea Drilling Project (DSDP) Leg 56/57 Site 436 (Figure 1) consists of basaltic crust overlain by ~100–160 m of Cretaceous to Oligocene chert, ~20–50 m of Eocene to early Miocene smectite-rich clay, and ~200–350 m of Miocene to Quaternary biogenic mud (Kodaira et al., 2020; Moore et al., 2015; Nakamura et al., 2013; Shipboard Scientific Party, 1980).

Flexure of the Pacific plate into the subduction zone generates bending-related normal faults that offset the oceanic crust and overlying sediments (e.g., Masson, 1991). Horsts and grabens bounded by these normal faults have widths of ~5–10 km and occur from the trench to ~120 km seaward of the trench axis. Fault offsets are ~100 m and increase towards the trench axis up to ~500 m (Boston et al., 2014). Horsts and grabens are also imaged beneath the upper plate up to 50 km landward of the trench axis, at depths up to 15 km, and with fault offsets up to 2 km (Tsuru et al., 2000; Kodaira et al., 2017). Some horsts and grabens on the incoming plate are associated with petit spot volcanism. Petit spot volcanoes are young (0–8 Ma) volcanic deposits, <300 m in height and <5 km in diameter, formed as a result of subduction-related plate flexure outboard of the trench axis (Hirano et al., 2006, 2019).

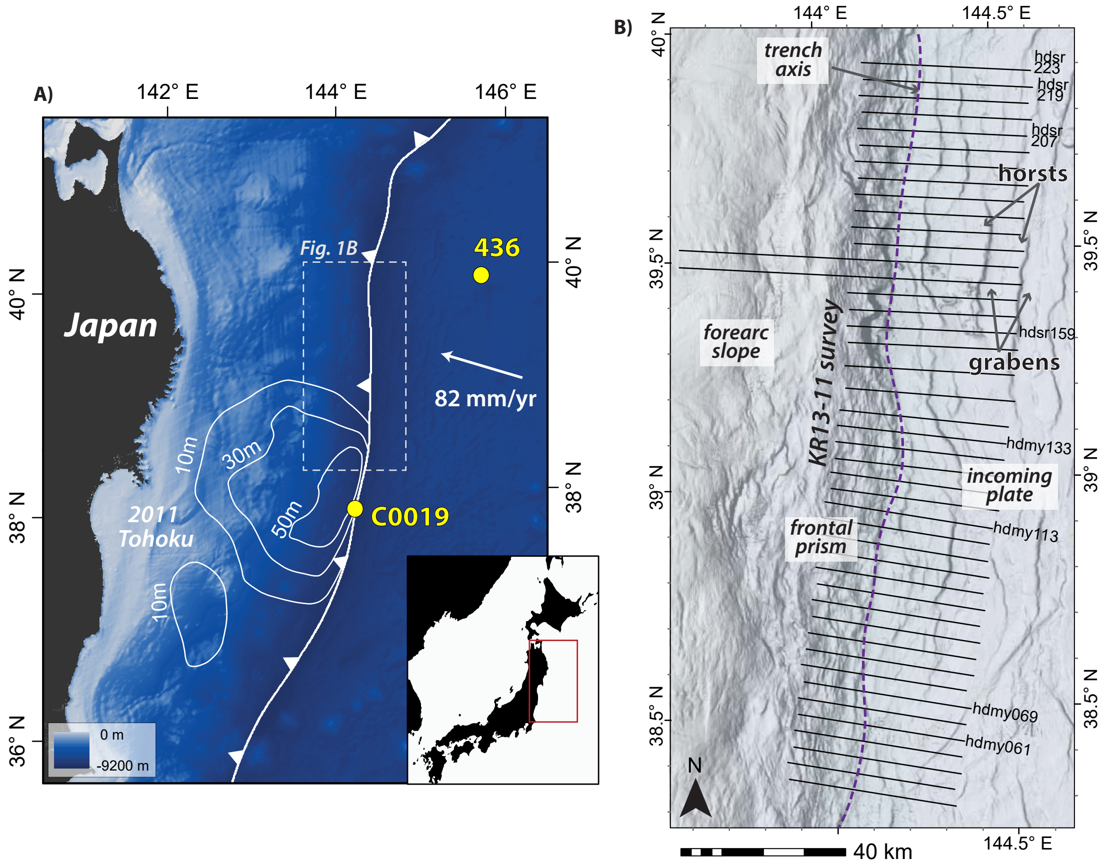


Figure 1 Location and tectonic setting of the study area at the Japan trench, offshore northeast Japan. **A.** Regional map showing the location of the trench, the plate convergence vector, and the M_W 9.0 2011 Tohoku earthquake rupture slip contours (white lines, after Iinuma et al., 2012). The study area (white dashed box, Figure 1b) is a ~40 km x 180 km region where 44 seismic reflection lines were collected by JAMSTEC (survey KR13-11). The survey area overlaps with the northern portion of the 2011 earthquake rupture. Locations of DSDP Site 436 and IODP Core C0019, which provide key stratigraphic data, are shown with yellow circles. **B.** Hillshaded 85 m DEM of the region in the white dashed box in panel A showing the trench axis (purple dashed line), forearc slope, frontal prism, and incoming plate horsts and grabens. Black lines show the locations of individual seismic reflection profiles mapped in this study.

3 Methods

We use 44 seismic reflection profiles that are 40 km in length and spaced ~4 km apart along the trench axis from ~38–40°N, collected and processed by the Japan Agency for Marine-Earth Science and Technology (JAMSTEC) (Figure 1). Lines were collected with Deep Sea Research Vessel *KAIREI* during the KR13-11 cruise using a cluster gun array with a volume of 380 in^3, 37.5 m shot interval, 6.25 m receiver interval, and a 1300 m-long streamer cable. The data are post-stack depth migrated lines. The velocity model used in the post-stack depth migration consists of three layers. The top layer is the water column and has a fixed velocity of 1525 m/s. The second layer is the soft sedimentary layer (SU1 and SU2). The velocity of the top of this layer is 1600 m/s and is linearly increased down to the bottom of this layer, with a 0.5 (m/s)/m gradient (namely, 1800 m/s at 400 m below seafloor, 2000 m/s at 800 m below seafloor). The model of the second layer was determined by comparing post-stack depth migrated profiles with several different gradient values. The gradient 0.5 (m/s)/m was chosen as it generated the best migrated image. The third layer is the chert and basement units (SU3 and SU4) and has a fixed velocity of 3000 m/s. We use these depth-migrated seismic reflection profiles to map seismic units (SUs) and faults, and to calculate sediment thickness and fault throw.

3.1 Seismic line mapping

We map seismic units on the incoming and overriding plate in the study area using the seismic stratigraphy of Nakamura et al. (2013, 2020, 2023) (Figure 2a). This work defines four seismic units (SU1 though SU4) based on seismic characteristics and correlations to known seismic and lithostratigraphic units of DSDP Site 436 and International Ocean Drilling Program (IODP) Expedition 343 Core C0019 (Nakamura et al., 2013). Seismic unit SU1, located landward of the trench, is an acoustically chaotic unit containing weak, discontinuous reflections. This unit is interpreted as the older, deformed equivalent of off-scraped incoming plate sediments (Nakamura et al., 2013, 2020; Regalla et al., 2019). Seismic unit SU2, located on the incoming plate and in frontal thrusts of the frontal prism, contains subhorizontal, parallel reflections with weak amplitudes. This unit is correlative to incoming plate sediments composed of diatomaceous muds with a basal pelagic clay layer (Nakamura et al., 2013). We include trench fill deposits, parallel and continuous reflections that overly the incoming plate sediments near the deformation front, as part of the SU2 map unit (Nakamura et al., 2023). Seismic unit SU3 contains high amplitude, sub-horizontal, semi-continuous reflections and is correlated to a chert layer stratigraphically below SU2. Seismic unit SU4 contains high amplitude reflections that become more discontinuous and weaker amplitude with depth. This unit is correlated to the basaltic crust of the incoming plate underlying SU3 (Figure 2a).

We identify faults in the incoming and overriding plate where there is folding, truncation, or offset of parallel reflectors in seismic units, or where high amplitude reflectors truncate at seismic unit boundaries or the seafloor. We map three categories of faults: imbricate thrust faults in the frontal prism, the shallow plate boundary décollement, and normal faults that offset seismic units in the incoming plate. Imbricate thrust faults are identified as low angle (<30°) thrust faults that sometimes appear as high amplitude reflections separating repeating packages of SU2 in the frontal prism (Figures 3a and b and S1). These imbricate thrust faults sole into the décollement or higher in the sediment section. The décollement is a high-amplitude, semi-continuous reflection that separates undeformed strata in the subducting plate from deformed material in the overriding plate (Nakamura et al., 2013, 2020). Prior mapping demonstrates that the décollement is localized within SU1 or within or at the base of SU2 and is generally parallel to the top of the SU3 and SU4 units in the subducting plate (Nakamura et al., 2013, 2020). Normal faults on the incoming plate and subducting plate beneath the prism are identified based on truncation of units SU2, SU3, and SU4 where they are offset at a high angle (>50°) by landward or seaward dipping faults (e.g., Boston et al., 2014; Nakamura et al., 2023).

We also map petit spot volcanic deposits on the incoming plate, slope failures at the toe of the frontal prism, and the position of the deformation front at the trench axis. Petit spot volcanic deposits are identified as mounds or knolls with discontinuous, high amplitude seismic reflections that disturb the top of the oceanic crust and sediment section (Fujiwara et al., 2007; Fujie et al., 2020). Slope failures are identified based on the presence of steep, arcuate topographic slopes and associated mass transport deposits (Nakamura et al., 2020). The deformation front is mapped at the position where deformed frontal prism sediments intersect the undeformed sediments of the incoming plate. We infer the position of the deformation front where slope failures modify the frontal prism.

3.2 Incoming plate sediment thickness and normal fault throw

We measured the thickness of incoming plate sediments at multiple locations on each seismic reflection profile (Figure 1) to determine the thickness of sediment available for accretion. Prior seismic mapping, core, and borehole data indicate that the biogenic muds and clays of seismic unit SU2 can be accreted to the upper plate, whereas the more strongly lithified cherts of seismic unit SU3 appear to evade accretion and are largely subducted (e.g., Nakamura et al., 2013, 2020; Chester et al., 2013). We therefore calculate the incoming plate sediment thickness within SU2 as a proxy for the total thickness of sediment available for accretion. Here, the term "incoming plate sediment thickness" refers to the biogenic muds and pelagic clays of SU2 and does not include the cherts of SU3.

We define the thickness of SU2 as the distance between the top of the upper most strong reflection of SU3 and the reflection at the top of SU2 at the seafloor (Figure 2b). We measure apparent sediment thickness as the vertical distance between these two reflectors, then convert to true sediment thickness using local strata dip. Measurement locations were selected at sites that minimized the impacts of local sediment re-distribution and alteration. Specifically, we chose locations that are >1 kilometer away from either outer-rise faults, which generate local sediment erosion and redistribution; or petit spot volcanism, that can locally alter sediment thickness or composition.

We quantified outer-rise fault offset by measuring the vertical separation of stratigraphic units across normal faults by projecting offset horizons to the fault midpoint (Figure 2c). We chose to measure vertical separation rather than fault throw because fault throw can only be accurately measured when there are discrete faults with distinct stratigraphic cutoffs imaged in the seismic reflection profile. However, steeply dipping fault planes and stratigraphic cutoffs are rarely clearly imaged in the seismic reflection data (e.g., Figures S3-8). In contrast, vertical separation of offset layers can be measured even if a fault plane or cutoff cannot be directly imaged, because the offset stratigraphic horizons can be projected into the fault zone (Figure 2c). In this study area, vertical separation and fault throw agree to within +/- 50 m (calculated for fault dips >50° and strata dips of <5°).

We calculated vertical separation of the top of unit SU3 because the cherts of SU3 were deposited prior to the onset of outer-rise faulting, and therefore record total fault slip. We measured vertical separation of SU3 at locations where the horizon is clearly imaged and where unit truncations have well constrained positions. These selection criteria allowed us to measure vertical separation at two to four faulted locations per profile on the incoming plate outboard of the trench, and up to three locations on the subducted plate under the frontal prism.

We ranked the confidence of each vertical separation measurement from 1 (low) to 3 (high) based on the distinctness of the fault plane and stratigraphic cutoffs. For example, sites with a high confidence include a single, discrete, fault with clear cutoffs, whereas sites that contain multiple faults or have unclear unit truncations have low confidence. We then calculated weighted averages of fault throw using the following equation:

$$W = \frac{\sum \int_{i=1}^{n} w_i X_i}{\sum \int_{i=1}^{n} w_i} \tag{1}$$

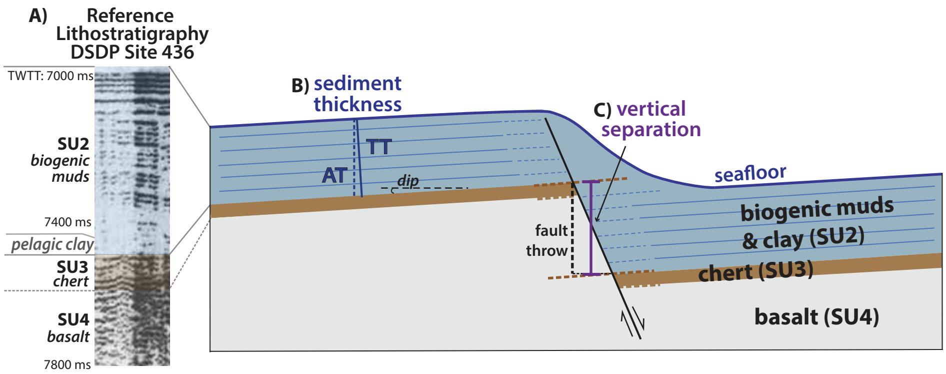


Figure 2 Schematic diagram of the incoming plate on a seismic line showing our approach to mapping seismic units (SUs), quantifying vertical separation (fault throw), and measuring sediment thickness. **A.** Seismic units (SUs) were mapped based on established correlations to known seismic and lithostratigraphic units of DSDP drill core at site 436 and to prior mapping of the margin (Nakamura et al., 2013, 2020; Nasu et al., 1980). See Section 3.1 for descriptions. **B.** Schematic diagram of a normal fault offsetting the incoming plate. True sediment thickness (TT) is calculated from apparent sediment thickness (AT), measured as the vertical depth difference between the seafloor reflection and the strong reflection of SU3 (chert), and the local dip of the strata. The dark blue line represents the seafloor reflection, which has the shape of a diffuse curve across the fault due to sediment redistribution and onlapping during fault slip. Reflectors within SU2 and SU3 are dashed where poorly resolved. **C.** Fault throw is determined by measuring the vertical separation of the chert horizon (SU3) across the fault by projecting the top of SU3 on the upthrown and downthrown blocks into the plane of the fault and calculating the vertical depth difference at the fault midpoint.

where W is the weighted average, w_i is site confidence weight, and X_i is site vertical separation. This approach allows us to identify spatial trends in normal fault throws that are most dependent on our highest confidence measurements.

4 Results

4.1 Seismic units and structure

4.1.1 Properties of the incoming plate outboard of the trench

Three seismic units are present on the incoming plate in the map area: SU2, SU3, and SU4. SU4 (basaltic crust) is mapped everywhere on the incoming plate outboard of the deformation front (Figures 3 and 4). SU3 and SU2 (chert, clay, and biogenic muds) are present nearly everywhere above SU4 on the incoming plate, except in regions where petit spot volcanic deposits are present. Petit spot volcanoes mapped in seismic reflection data are ~1–5 km in length and ~100–500 m in height (Figure 3e). There are three main clusters of petit spot volcanic deposits located at 39.25°N–39.75°N, 38.7°N–38.9°N, and possibly at 38.4°N (Figure 5).

Incoming plate sediment thickness varies between ~50 and 400 m, with more than half of the measured sites having thicknesses of 250–400 m (Figure 5a). There are, however, two ~20 x 20 km regions of anomalously thin (~50–200 m) sediments located at ~39.5°N and ~38.75°N. These regions are located near mapped locations of petit spot volcanism, where SU2 is thin or absent (e.g., Figure 3e). This spatial correlation between thin sediments and petit spot volcanism has been suggested to be a result of volcanic intrusions that metamorphose and effectively thin the muds and clays of the sediment section (e.g., Fujie et al., 2020; Hirano et al., 2006). The thin sediment regions we map are surrounded by ~10 km wide margins with intermediate sediment thicknesses (~200–300 m), and transition into "background" sediment thicknesses (250–400 m) over distances of ~>50 km.

Outer-rise faults offset seismic units SU2, SU3, and SU4 throughout the study area (Figures 4 and 5b). Normal faults outboard of the trench have strikes that are subparallel to the trench and are present from 0 to >100 km. Both landward and seaward dipping normal faults are present and bound networks of horsts and grabens. Normal faults have throws of ~90 m to 900 m, and generally increase in magnitude with proximity to the trench, in agreement with the findings of Boston et al. (2014). We note that there is one 20 km long region on the incoming plate at 39.4°N that contains a cluster of normal faults with very high fault throw (>300 m). This region is located in the vicinity of the petit spot volcanism around ~39.5°N (Figure 5b).

4.1.2 Properties of the subducted plate under the frontal prism

Seismic units SU4, SU3, and SU2 are also present under the frontal prism. SU4 (basaltic crust) is mapped everywhere on the subducted plate beneath the frontal prism (Figures 3 and 4). SU3 can be only definitively mapped under the frontal prism to within <~5 km of the deformation front, at which point the resolution of the seismic data often precludes separation of SU4 from SU3 (Figure 4). SU2 is typically not present on the subducting plate under the frontal prism, except at limited locations where it is present in subducted grabens (Figure 3b and d) or is underthrust beneath the frontal prism (Figure 3c).

We mapped subducted normal faults landward of the

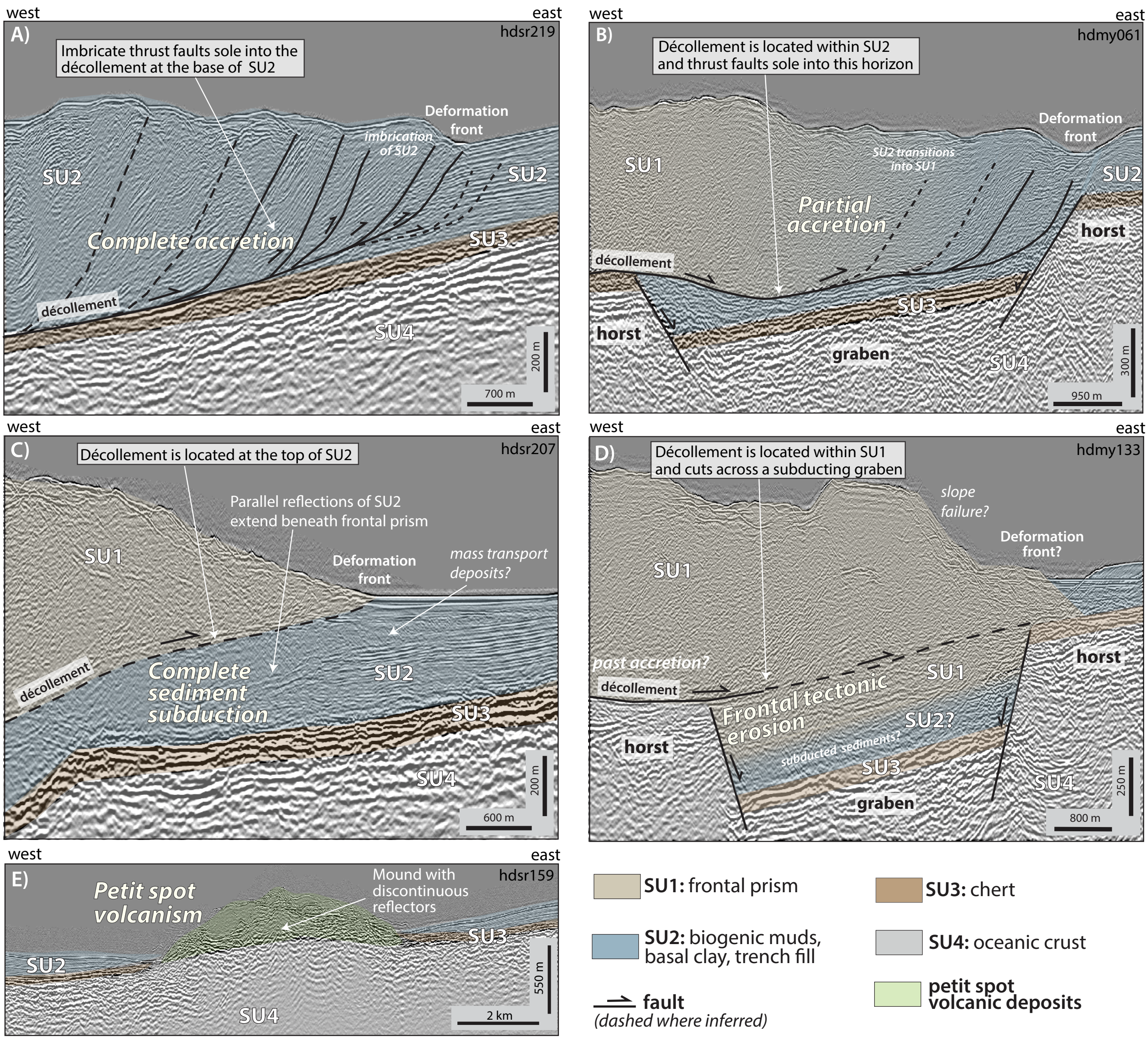


Figure 3 Annotated seismic lines showing examples of mapped seismic units, normal faults, imbricate thrust faults, and décollement position (see Supporting Information Figure S1 for non-annotated images). **A.** Example of a location where imbricate thrust faults in SU2 sole into the décollement at the base of SU2 (biogenic mud). We interpret this geometry to indicate that all SU2 sediments are accreted at the deformation front and SU3 and SU4 are locally subducted. **B.** Example of a location where imbricate thrust faults sole into décollement within SU2, rather than at its base. We interpret this geometry to represent partial sediment accretion, where SU2 above the décollement is accreted and SU2, SU3, and SU4 below are subducted. **C.** Example showing the décollement at the top of SU2 such that there is complete sediment subduction of units SU2, SU3, and SU4 beneath the frontal prism (SU1). **D.** Example of a location where the décollement appears to cut across SU1 to connect the tops of two horsts. Also note the steeply dipping slope at the deformation indicative of slope failure. We interpret this décollement geometry to indicate that part of SU1, and likely all of SU2, SU3, and SU4, are locally subducted. This geometry may lead to frontal tectonic erosion of prism sediments (e.g., Hilde, 1983). **E.** Example of petit spot volcanism on the incoming pate. Petit spot volcanoes are identified as seismically chaotic mounds (light green) that intrude and disrupt the oceanic crust and sediment section.

deformation front (under the frontal prism) and interpolated fault boundaries between adjacent seismic lines (Figure 5b). Normal faults subducted beneath the frontal prism branch and merge along strike and bound horsts and grabens that have length and widths comparable to that on incoming plate (Figure 5b and c). In some locations, the displacement along normal faults decreases along strike such that there are neither horsts nor grabens present in a line segment. These likely represent regions where there is flat subducting topography or relay ramps between adjacent faults. The average throw on normal faults under the frontal prism is 354 ± 163 m. This magnitude of offset is greater than the average throw on outer-rise faults outboard of the trench (261 ± 103 m), and implies that there is a continued accumulation of slip on normal faults following

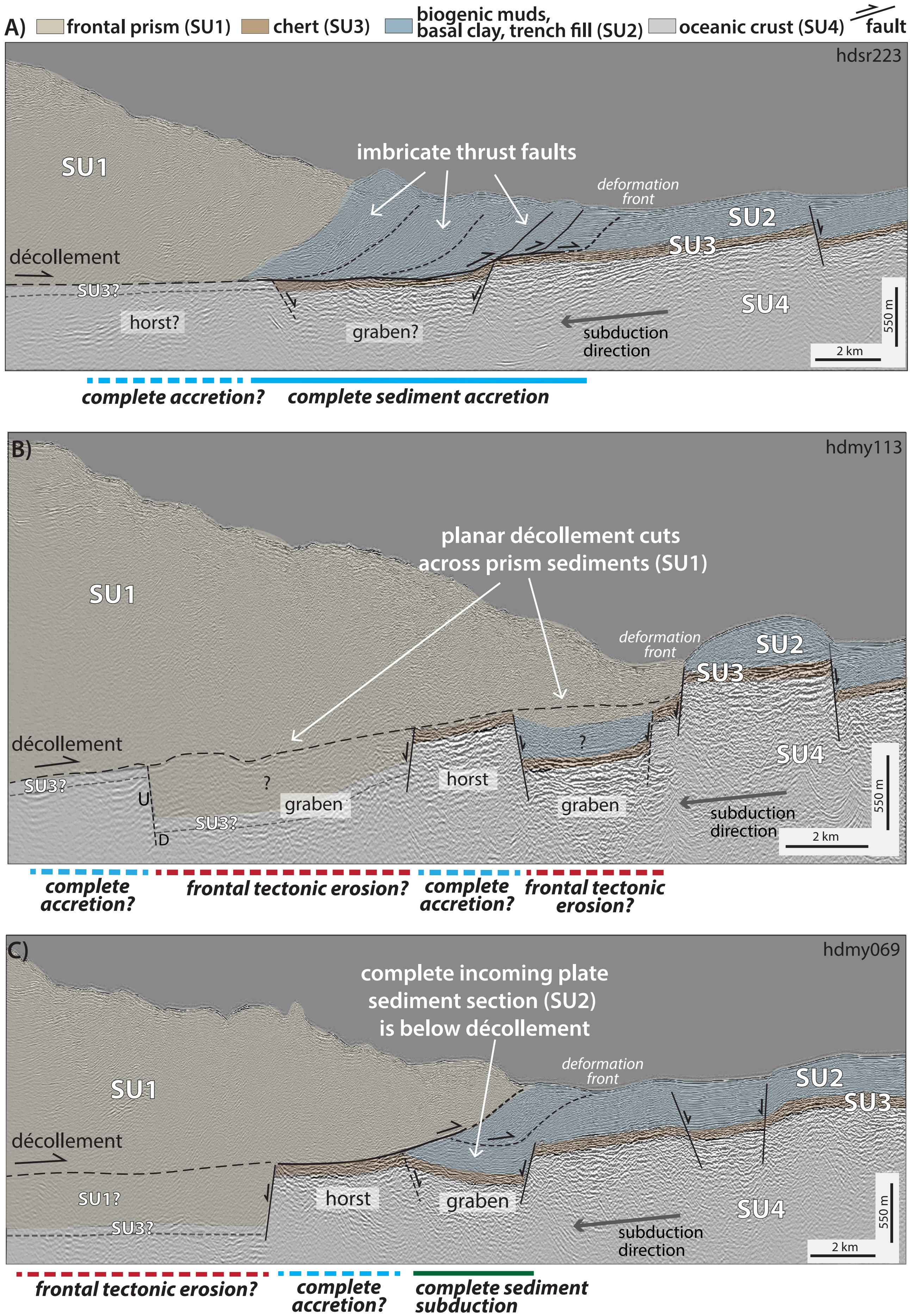


Figure 4 Caption on next page.

Figure 4 Examples of variations in mapped deformation modes, sediment flux, and décollement position within a seismic reflection profile. See Figures 1 and 5 for profile locations and the Supporting Information figures S3 – 5 for un-annotated versions of these seismic profiles. Inferred position of SU3 is shown as dashed lined on the landward portions of the profiles. **A.** Line hdsr223 has a single mode of deformation along the entire seismic profile. There is complete sediment accretion near the deformation front and inferred sediment accretion along landward segments. **B.** Line hdmy113 has two modes of deformation, alternating between frontal tectonic erosion within subducted grabens and sediment accretion on horsts. **C.** Line hdmy069 has three modes of deformation: complete sediment subduction near the deformation front, possible sediment accretion on the horst top, and potential tectonic erosion in the landward-most half-graben.

subduction under the frontal prism.

4.1.3 Properties of the overriding plate

Two seismic units are present in the overriding plate: SU1 (chaotic frontal prism) and deformed SU2 (biogenic muds and clay) (Figures 3 and 4). Unlike many margins with well-formed accretionary prisms, at the Japan trench it is difficult to resolve coherent reflectors within the frontal prism or identify internal deformation structures within SU1 (Nakamura et al., 2013, 2020). In contrast, where SU2 is present in the overriding plate, coherent reflectors are deformed into fault-related folds as the incoming sediments are repeated in imbricate thrusts that sole into a local décollement (Figure 3a and b). Imbricate thrust faults repeating SU2 are present in ~2/3 of the mapped seismic lines and comprise the outermost portion of the frontal prism (Figures 3a and b; 4a and c). In the other ~1/3 of the lines, SU1 is present all the way to the deformation front (Figures 3c and d; 4b). The nature of the boundary between SU1 and SU2 varies among seismic lines and can either be gradational (e.g., Figure 3b) or abrupt (e.g., Figure 4a).

In some locations, the internal structure of the prism is disrupted by slope failures that have heights of ~500 m–1 km and headwall slopes of ~20–30° (Figure 3d). The highest density of large slope failures occurs in a region between 39.3°N and 39.6°N (Figure 5c), inboard of a region on the incoming plate with very large offset outer-rise faults, thin sediments, and petit spot volcanism. This correlation suggests that slope failures may be promoted by the subduction of crust with large fault throw, thin sediment, and seafloor roughness generated by volcanism.

4.1.4 Properties of the décollement

The décollement in our study area is expressed in three ways: 1) a subhorizontal fault into which imbricate thrusts sole, 2) a high amplitude, subhorizontal reflector below SU1 and/or SU2 (Nakamura et al., 2013, 2020), or 3) the contact between SU1/SU2 and the underthrust units (Figure 3a-d). In locations where imbricate thrust faults sole into the base of SU2 at its contact with SU3 (Figure 3a), the décollement is likely hosted within a layer of smectite rich, frictionally weak clays (e.g., Chester et al., 2013; Kirkpatrick et al., 2015; Moore et al., 2015; Nakamura et al., 2013). Where the décollement is located within SU2 (Figure 3b), it is likely hosted within frictionally stronger biogenic muds (Ikari et al., 2015; Sawai et al., 2017).

High amplitude décollement reflectors mapped at depth may project into imbricate faults within the lower portion of SU2 (Figure 3a and b), or may project above SU3 and SU2, below SU1 (Figure 3c and d). Mapped décollement reflectors have two endmember geometries: a semi-planar trajectory across both horsts and graben (Figure 3d), or a non-planar geometry where it either steps down from a horst into a graben or steps up from graben into a horst (Figure 3b). The strength of the décollement reflector often changes as the décollement crosses into different seismic units. For example, a planar décollement that crosses from a horst top into prism sediment above a graben weakens in seismic character, whereas the décollement often maintains its seismic character as it steps down into a subducting graben (Figures 3 and S3-5).

In locations where there are neither imbricate thrust faults nor a clear high amplitude reflector, we infer the location of the décollement to be the position that separates SU1/SU2 material in the frontal prism from material on the subducted plate (Figure 3c and d), and it is inferred to have a position that is congruent with the décollement as mapped in the up-dip and down-dip directions. The unit or units present under the décollement vary as a function of distance from the trench, and as a function of its position relative to a subducted horst or graben. SU2 is mapped under the décollement in lines where some or all of the SU2 section is underthrust at the deformation front (Figure 3b and c) or at locations where SU2 is entrapped within a subducting graben (Figure 4c). Where SU2 is present under the décollement, it is usually undeformed and contains flat parallel reflectors (Figure 3b and c). Greater than ~1–3 km from the deformation front, the décollement typically separates SU1 in the overriding plate from SU3 or SU4 in the downgoing plate (Figure 4b and c), and SU2 is absent. Finally, in lines where deep (greater than ~200m) grabens are subducted, SU1 may be present beneath the décollement (Figures 3d and 4b), along with SU2, SU3, and SU4.

4.2 Frontal prism structure and fate of sediments

We combine our mapping of seismic units, imbricate thrust faults, and décollement position to generate criteria for identifying portions of the forearc that are the product of sediment accretion, sediment subduction, or frontal tectonic erosion (deformation "modes", Table 1). Because of the time transgressive nature of deformation during subduction, structures mapped at greater distances from the trench reflect the cumulative processes

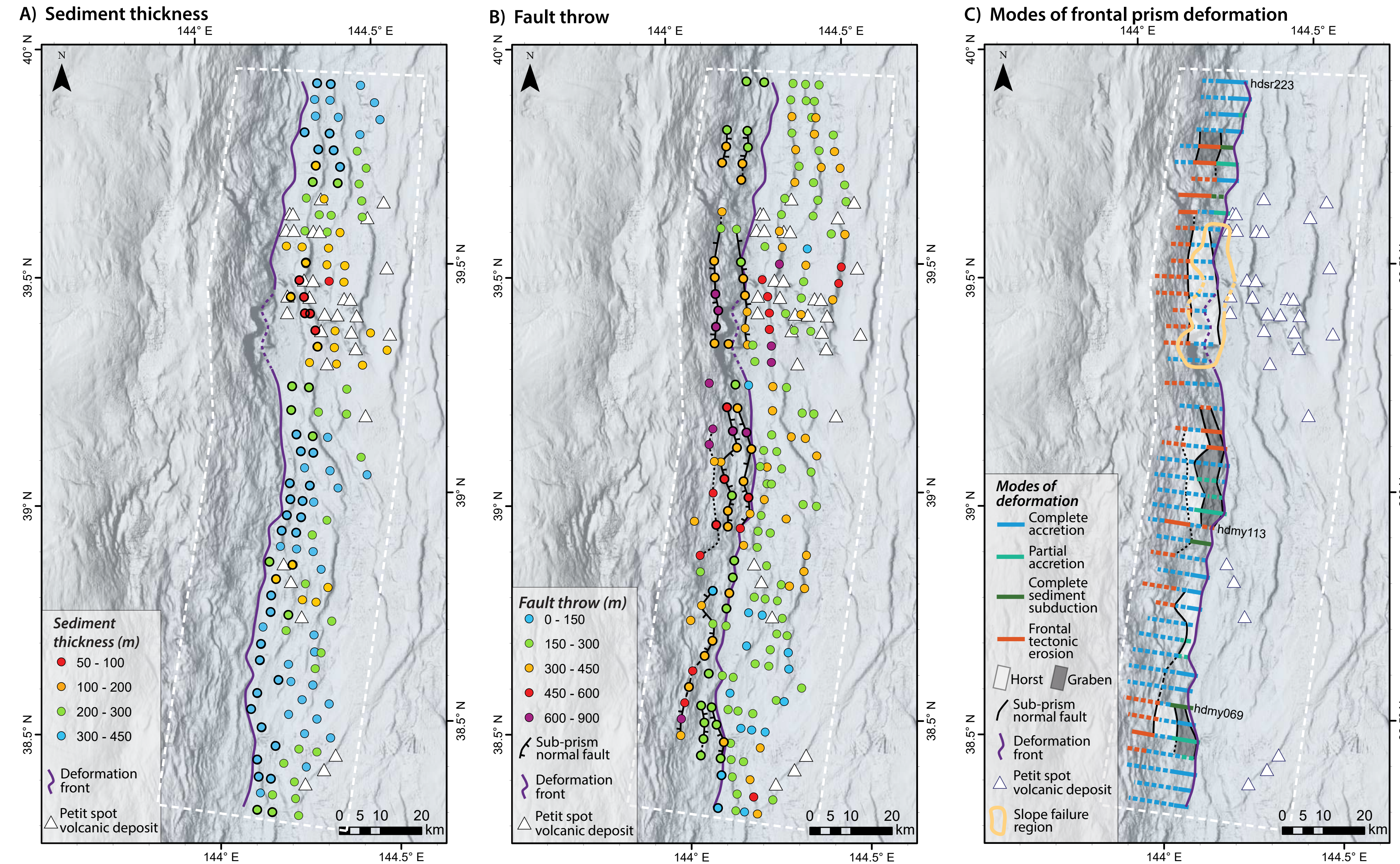


Figure 5 Caption on page 11.

that operated when that portion of the subducted plate was at the trench, while it was being translated to its current position, and those operating *in situ* today. As such, we define separate criteria to interpret each mode in the up-dip region at the deformation front, and down-dip regions ~>5km landward of the deformation front.

4.2.1 Complete sediment accretion

Near the deformation front, line segments are categorized as experiencing complete sediment accretion if they contain imbricate thrust faults in the frontal prism that sole into a décollement localized at the base of the incoming plate sediment section, SU2 (Table 1, Figure 3a). In these locations, SU3 and SU4 are located below the décollement. We interpret this geometry to indicate that the complete sediment section of SU2 is off-scraped and actively accreted to the frontal prism. Landward of the deformation front, we categorized a line segment as having experienced complete sediment accretion if the décollement separates prism sediments (SU1/SU2) in the overriding plate from the chert and basalt (SU3/4) of the subducting plate, with no incoming plate sediment (SU2) observed beneath the décollement. In these locations, we infer that the missing SU2 was previously accreted at the deformation front, and that the frontal prism is actively sliding along a décollement that is localized directly above chert/basalt of SU3/4.

4.2.2 Partial sediment accretion

Near the deformation front, line segments are categorized as experiencing partial sediment accretion if they contain thrust faults that sole into a décollement formed in a horizon within SU2 (Table 1, Figure 3b). In these locations, parallel reflections of SU2 are often observed beneath the décollement. We interpret that in these lines, the upper portion of the sediments above the décollement are actively accreting to the frontal prism and the sediments beneath the décollement are locally subducted. Landward of the deformation front, partial accretion is mapped in locations where the décollement separates prism sediments (SU1/SU2) in the overriding plate from incoming plate sediments (SU2) below the décollement. In these locations, either parallel reflectors of the lower part of SU2 are observed below the décollement, or the total thickness of SU2 under the décollement is much less than the total thickness of SU2 on the incoming plate. We infer that the missing upper section of SU2 was accreted at the deformation front, such that only the lower portion was subducted to depth. These data imply that the frontal prism is actively sliding along the décollement over the lower portion of SU2, such that there is active subduction of the lowermost incoming plate sediments.

4.2.3 Complete sediment subduction

Near the deformation front, line segments are categorized as experiencing complete sediment subduction if the décollement is localized at the top of the incoming plate sediments below prism sediments (Table 1, Figure 3c). In these locations, seismic reflectors that represent the complete thickness of the incoming plate sediments (SU2) are present beneath the décollement. Here, the complete sediment section is locally subducted and there is no active accretion occurring at the deformation front. Landward of the deformation front, complete sediment subduction is mapped using similar criteria: reflectors comprising the complete SU2 section are observed beneath the décollement. In these locations, we infer that no sediment accretion occurred at the deformation front, and that there is currently stable sliding of the prism along the décollement above the subducted sediment section.

4.2.4 Incipient frontal tectonic erosion

Frontal tectonic erosion, whereby frontal prism material is subducted to depth, has been proposed to occur by entrapment of frontal prism material under the décollement within a subducted graben (e.g., Hilde, 1983). This process has been proposed to initiate via gravitational collapse and slumping of frontal prism material that fills the incoming plate graben at the trench (e.g., Hilde, 1983; von Huene and Culotta, 1989). The slope failures and sediment slumping that we map at the trench may indicate the incipient process that can ultimately lead to frontal tectonic erosion (Figures 3d and 5c). Therefore, while we do not map frontal tectonic erosion as an active process at the deformation front, slope failures mapped at the deformation front may indicate locations where frontal tectonic erosion will occur in the future.

We can, however, map portions of line segments at depth that may be experiencing active frontal tectonic erosion. Frontal tectonic erosion is mapped at locations landward of the deformation front where a planar décollement cuts across a graben within prism sediments (SU1) such that frontal prism material and incoming plate sediments (SU2) are entrapped below the décollement within a graben (Figures 3d and 4b and c). In these locations, we infer that both tectonic erosion and subduction of incoming plate sediments is occurring. Frontal tectonic erosion is also inferred in locations where a décollement reflector is not clear, but where high relief subducted grabens are filled with SU1 and SU2. Where frontal tectonic erosion is mapped, we infer that the prism slides along the décollement and that both SU1 and SU2–4 are actively subducted.

4.3 Spatial variability in modes of deformation

Using the above criteria, we mapped modes of deformation on each seismic line, covering a region that extends ~180 km along margin strike, from the deformation front to ~20 km down dip (Figures 4 and 5c; S3–8). Our mapping shows that modes of deformation are not constant with depth along a single seismic line (Figure 5c), and only ~30% of the seismic profiles contain a single mode of deformation (e.g., Figure 4a). Most seismic lines contain two or three different deformation modes (Figures 4b and c; S4 - 8). Transitions between modes occur at horst or graben boundaries. For

Figure 5 Results of mapping seismic reflections profiles along the Japan trench. The white dashed boxes denote the extent of the seismic profiles in the study area. The purple line indicates the mapped location of the deformation front. The white triangles represent petit spot volcanic deposits on the incoming plate. **A.** Sediment thickness measurements at 130 locations on the incoming plate. **B.** Map showing measured throw along normal faults developed on the Pacific plate crust at 199 locations. Note that regions with petit spot volcanism are located with thin sediments and high offset faults. Thick outlined points in panels **A** and **B** denote sites used in ratio calculations at the trench. Only sites with both sediment thickness data within 10 km of the trench and fault throw data on the first subducted horst or graben were used. **C.** Map of modes of frontal prism deformation: solid line where certain; dashed where uncertain or inferred. These data show that prism evolution varies between sediment accretion, sediment subduction, and frontal tectonic erosion over ~5 - 10 km length scales along strike and down dip. Cross-comparisons of all three maps show that sediment accretion occurs in regions with thick incoming plate sediment and small fault throw, while slope failure and tectonic erosion tend to occur in places with thin sediments, petit spot volcanism, and large fault throw.

example, line hdmy113 (Figure 4b) is characterized by alternating modes of tectonic erosion in grabens and sediment accretion on horsts. Similarly, line hdmy069 (Figure 4c) is characterized by complete sediment subduction in a graben at the trench, complete accretion in the down-dip horst, and tectonic erosion in the next graben down dip. Our seismic line mapping shows that the mode of deformation varies in the study area over length scales of ~5–20 km along strike, and ~5–10 km down dip (Figure 5c).

Importantly, we find that the types of deformation that can occur, and therefore the possible pathways for incoming and upper plate sediments, differ between horsts and grabens. Horsts in the study area only experience complete sediment accretion or partial accretion, and do not experience tectonic erosion or sediment subduction (Figures 3, 4, and 5c; S3 - S8). This observation, apparent both at the trench and at depth, indicates that most or all of the sediment on subducted horsts is off-scraped by imbricate thrusts and accreted to the frontal prism at the deformation front.

In contrast, subducted grabens can experience any of the four modes of deformation (Figure 5c). Our mapping shows approximately half of the grabens experience complete or partial accretion and half experience complete sediment subduction or tectonic erosion. This observation of sediment accretion in grabens directly contrasts prior models that suggest subducted grabens ubiquitously lead to sediment subduction and promote frontal tectonic erosion (e.g., Hilde, 1983; von Huene and Culotta, 1989). Instead, our data imply that other processes may control whether sediments in a subducted graben will be accreted, subducted, or tectonically eroded.

4.4 Correlations between sediment thickness, fault throw, and frontal prism deformation

The relative magnitudes of outer-rise fault throw and incoming plate sediment thickness at the time of subduction may play a key role in sediment flux and prism evolution. We test this hypothesis by correlating the maps of frontal prism deformation mode with measurements of fault offset and sediment thickness on the incoming plate at the trench (Figure 6). Along each seismic line, we calculate the weighted-average fault throw from the first one to two normal faults subducted beneath the frontal prism and average sediment thickness using measurements within ~10km of the trench (See Figure 4 for measurement locations). We then correlate these near-trench values with the mode of deformation mapped closest to the trench in order to determine variations in sediment thickness and fault throw as a function of deformation mode.

We find that incoming plate sediment thicknesses are similar (~300–365 m) in regions experiencing complete accretion, partial accretion, complete sediment subduction, and frontal tectonic erosion, whereas sediment thickness outboard of regions with slope failures are exceptionally low (115 ± 45 m) (Figure 6a). We find that fault throw varies between the different modes of deformation, where larger fault throw values correlate to smaller volumes of accreted sediment. Complete sediment accretion occurs where the subducting oceanic crust has the smallest average fault throws of 216 ± 65 m. Partial accretion and complete sediment subduction occur where the subducting crust has intermediate fault throws of ~295–340 m (Figure 6b). Portions of the margin experiencing frontal tectonic erosion and slope failure have significantly larger fault throws, averaging around ~490 - 500 m.

We evaluate this relationship between the incoming plate and frontal prism deformation by combining sediment thickness and fault throw measurements (Figure 6a-b) into a single ratio using near-trench measurements and correlate these near-trench ratios to the mode of deformation mapped closest to the trench (Figure 6c). These ratio data show discrete variations between the relative magnitudes of sediment thickness and fault throw as a function of deformation mode. Complete sediment accretion occurs inboard of regions where there is a high ratio of sediment thickness to fault throw (1.50 ± 0.6) (Figure 6c). Partial sediment accretion and complete sediment subduction occur inboard of regions where the ratio of sediment thickness to fault throw on the incoming plate have intermediate values of 1.09 ± 0.2 and 1.25 ± 0.2 respectively. Frontal tectonic erosion occurs inboard of regions with a low ratio of sediment thickness to fault throw (0.72 ± 0.2). Lastly, slope failures at the deformation front occur where the lowest ratio of sediment thickness to fault throw (0.28 ± 0.2).

We note that the total number of observations for some modes of deformation is small (n=2 to 6). If we

Mode of deformation	At deformation front		Landward of deformation front	
	Mapping criteria (observations)	Interpreted processes	Mapping criteria (observations)	Interpreted processes
Complete sediment accretion	• Imbricate thrust faults sole into the décollement • Décollement is localized at the base of the incoming plate sediment section (SU2) • SU3 and SU4 are below the décollement	• Complete incoming sediment section is actively accreting to the frontal prism	• Décollement is localized above chert/basalt (SU3/SU4) • No incoming plate sediment (SU2) is observed beneath the décollement	• Prism (SU1) is sliding along décollement • Incoming plate sediment (SU2) was previously accreted
Partial sediment accretion	• Imbricate thrust faults sole into the décollement • Décollement is localized within the incoming plate sediment section (SU2) • Parallel reflections of lower SU2 are present beneath décollement	• Upper portion of incoming plate sediments (SU2) is actively accreting • Lower portion of SU2 is locally subducted	• Décollement separates prism sediments (SU1/SU2) from lower portion of incoming plate sediments (SU2) • Parallel reflectors of SU2 observed below décollement	• Prism (SU1) is sliding along décollement • Upper portion of SU2 was previously accreted • Lower portion of SU2 is locally subducted
Complete sediment subduction	• Décollement is localized between the prism sediments (SU1) and the top of the incoming plate sediments (SU2) • Complete set of SU2 reflections is present beneath the décollement	• Complete sediment section is locally subducted; no accretion occurring	• Décollement is localized at the top of the incoming plate sediment section (SU2) • Parallel reflectors (SU2) are present beneath décollement • Décollement separates prism sediments (SU1) and incoming plate sediments (SU2)	• Prism (SU1) is sliding on top of subducted sediment section (SU2) • Sediments below décollement are being subducted • No accretion has occurred
Slope failures and incipient frontal tectonic erosion	• *Frontal tectonic erosion does not occur at the deformation front; it only occurs landward of the deformation front*	• *The precursor to frontal tectonic erosion may be slope failures at the deformation front and mass transport deposits into a subducting graben*	• Planar décollement cuts across a graben within prism sediments (SU1), connecting two adjacent horsts • Frontal prism material (SU1) and incoming plate sediments (SU2) are entrapped below the décollement within a graben	• Stable sliding across décollement • Frontal prism sediments (SU1) and incoming plate sediments (SU2) within graben are locally subducted

Table 1 Criteria for mapping modes of frontal prism deformation.

increase the number of observations by correlating all measurements of fault throw and sediment thickness on the incoming and subducted plates with each type of deformation mode at the trench, we see similar relative relationships between fault throw, sediment thickness, and deformation mode. However, these data show greater overlap, with ratios of 1.15 ± 0.6 for sediment accretion, 1.15 ± 0.4 for partial accretion, 1.06 ± 0.4 for complete sediment subduction, 0.94 ± 0.5 for frontal tectonic erosion, and 0.41 ± 0.2 m for slope failure (Figure S9). Similarity in these values may be the result of averaging across variable fault throws and sediment thicknesses that exist over larger areas (~15–20 km) as compared to near trench data (5–10 km) (Figures 5 and 6d). This suggests near-trench values of sediment thickness and fault throw are the better predictor of deformation mode at the trench. These data also suggest that small-spatial scale variation in fault throw or sediment thickness on the incoming plate can lead to significant spatio-temporal variation in deformation mode at the trench.

5 Discussion

5.1 Influence of sediment thickness and fault throw on modes of frontal prism deformation and sediment flux

Modes of frontal prism deformation and the fate of both incoming and upper plate sediments are a direct function of whether horsts or grabens are being subducted and the relative magnitudes of fault offset versus sediment thickness. First, we find that sediment accretion always occurs on horsts, but grabens experience all modes of deformation. Second, we find that sediment accretion at the deformation front is promoted where there are thicker incoming plate sediments (~300 m) and outer-rise faults with low offsets (~225 m); that is, in regions where the ratio of sediment thickness to fault throw is >1 (Figures 5 and 6). Third, we find that locations on the incoming plate where there is petit spot volcanism, moderate or thin sediments (<~300 m), and outer-rise faults with large offsets (>350 m) correlate to locations on the overriding plate with slope failures and frontal tectonic erosion, where the ratio is <1 (Figures 5 and 6). These observations suggest that there may be a critical threshold of sediment thickness to fault throw that controls the modes of deformation and sediment flux in the presence of subducting grabens. At the Japan trench, this threshold appears to be approximately 1:1 ratio of sediment thickness to normal fault throw.

5.2 Implications for décollement heterogeneity

Spatio-temporal variations in the ratios of incoming plate sediment thickness to outer-rise fault throw, and thus modes of frontal prism deformation, directly impact décollement geometry and mechanics by dictating the stratigraphic position of the décollement and by introducing geometric barriers (horsts and grabens) to the décollement. We observe spatial variations over short length-scales (~5–10 km) in both frontal prism mode of deformation and incoming plate ratio of sediment thickness to fault throw (Figures 5 and 6d). Therefore, equivalent heterogeneity in décollement geometry and composition must occur at similar spatial scales.

Our data suggest that there may be a threshold of fault throw, given an incoming sediment thickness, that if exceeded, disrupts the décollement's ability to localize at the base of the incoming sediment section. For normal fault throws below this threshold (ratio >1), a single continuous, undulating décollement can develop that promotes sediment accretion (Figure 7a). For fault throws above this threshold (ratio <1), a planar décollement develops across a graben that is mechanically more favorable to propagate that promotes sediment subduction and frontal tectonic erosion (Figure 7b).

Variations in the occurrence of sediment accretion, sediment subduction, and frontal tectonic erosion therefore have direct implications for two types of heterogeneity on the shallow décollement: geometric heterogeneity and compositional heterogeneity. First, subducting horsts and grabens on the incoming plate modulate shallow décollement mechanics by creating geometric barriers over which the décollement must propagate. We observe two different endmember geometries of the décollement where it interacts with horst and graben topography. At the Japan trench, when sediment thickness is large relative to fault throw, we observe that the décollement undulates with subducting topography, stepping up and down subducted normal faults to maintain its position in a similar stratigraphic horizon (e.g., Figures 4a and 7a). This stepping, however, creates geometric barriers over which the prism must slide, and these barriers may create local stress heterogeneities along the shallow megathrust (e.g., Sun et al., 2020). In contrast, when sediment thickness is small relative to fault throw, we observe that the décollement does not step down into the graben, but instead attempts to maintain its planarity as it propagates across the adjacent graben (e.g., Figures 4b and 7b). This planar décollement "smooths" over the subducting topographic roughness generated by normal faults without developing normal fault-related geometric barriers in the shallow megathrust. Therefore, the ratio of sediment thickness to fault offset on the incoming plate may be used to infer the degree of geometric heterogeneity present in the shallow décollement, where high ratios predict undulating décollements that step over geometric barriers, and low ratios predict planar décollements that smooth over these barriers.

Second, spatial heterogeneity in the mode of frontal prism deformation requires correlative heterogeneity in the composition of the shallow décollement. This is because in systems with subducting outer-rise normal faults, the juxtaposition of upthrown and downthrown horst and graben blocks on the incoming plate creates lateral variations in the incoming plate sediment section. A décollement that has a planar geometry that cuts across subducting grabens will inherently propagate across different sedimentary units. In contrast, a décollement that undulates with the subducting topography may remain in a continuous stratigraphic horizon. In the Japan trench, where sediment thickness is large relative to fault throw, we observe that the décollement undulates with subducting horsts and graben topography in order to remain in the same stratigraphic position near the SU3 - SU2 contact (Figure 7a). Conversely, where sediment thickness is small relative to fault throw, the décollement is planar and may cut across different lithologic units, including SU2 and SU1, introducing compositional heterogeneity to the décollement (Figure 7b). Therefore, the ratio of sediment thickness to fault throw on the subducting plate may be a predictor for compositional heterogeneity on the shallow plate interface, where high ratios correlate to compositionally homogenous décollement segments and low ratios correlate to heterogeneous décollement segments.

These observations imply that, in portions of the Japan trench where sediment thickness is greater than normal fault throw, it is mechanically more favorable for the décollement to make a small bend and remain in a continuous stratigraphic horizon than to develop a new décollement segment that smooths over subduct-

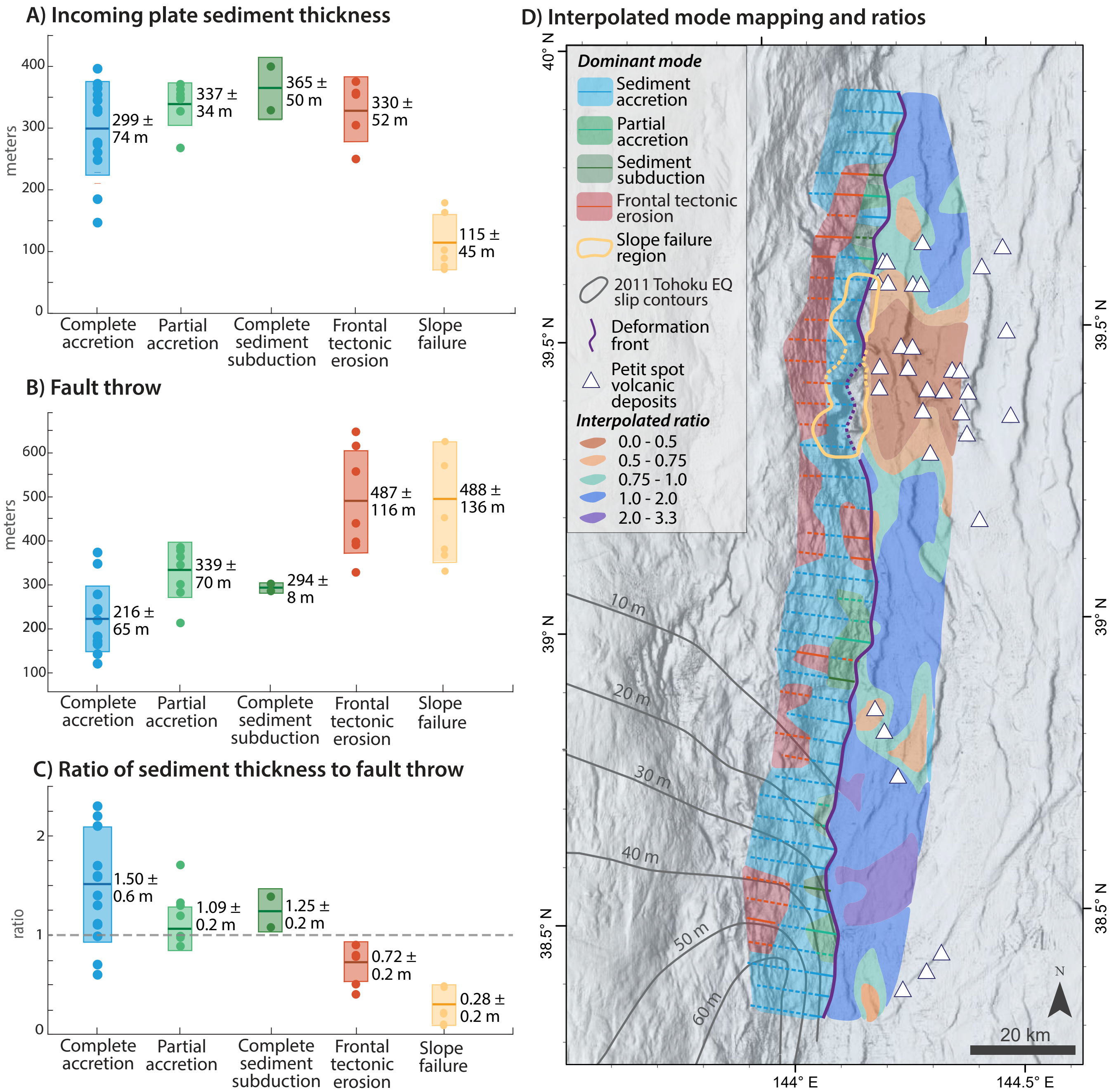


Figure 6 Comparisons between sediment thickness, fault throw, and mode of frontal prism deformation at the trench. Box plots **(A - C)** include individual measurements (points), weighted averages (horizontal line), and $2\sigma_w$ uncertainty (boxes). See Figure 4 for measurement locations. **A.** Sediment thickness as a function deformation mode. Average sediment thickness is the smallest outboard of regions with slope failures. Average sediment thickness is similar in regions experiencing all other modes. **B.** Fault throw as a function of deformation mode. Portions of the prism experiencing complete accretion have the smallest average fault throw, followed by partial accretion and full sediment subduction, frontal tectonic erosion, and regions with slope failures. **C.** Ratio of sediment thickness to fault throw as a function of deformation mode. Regions experiencing complete accretion, partial accretion, and complete sediment subduction have ratios >1. Regions experiencing tectonic erosion or slope failure have ratios <1. **D.** Map of dominant prism deformation mode, interpolated sediment thickness to fault throw ratio on the incoming plate (1km grid), and slip contours for the 2011 Tohoku earthquake (after Iinuma et al., 2012).

ing topography. Only when normal fault throw is large is it more favorable to develop a new planar décollement segment that propagates through subducting sediments or frontal prism. This mechanical favorability may be influenced, at least in part, by the frictional properties of the sediments in the incoming section. Coring of the incoming plate, frontal prism, and décollement at IODP site C0019 and DSDP site 436 show that the décollement is locally developed within the basal, smectite rich clay layer, correlative to the basal portion of SU2 on the incoming plate (Chester et al., 2013; Nakamura et al., 2013; Kirkpatrick et al., 2015; Nasu et al., 1980). Friction experiments indicate that these basal clays have high concentrations of smectite and lower friction co-

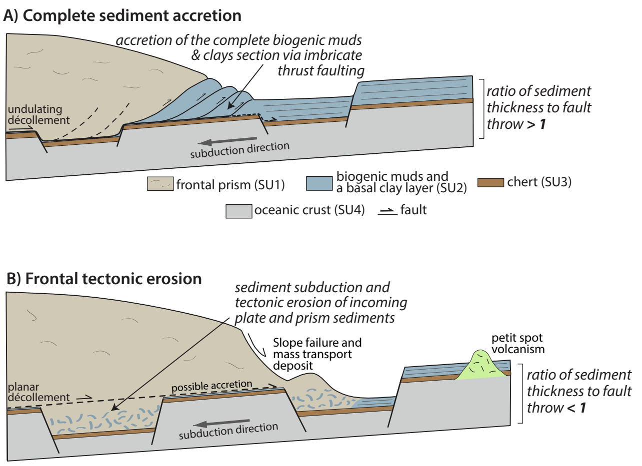


Figure 7 Endmember models of modes of frontal prism deformation based on a synthesis of mapped seismic reflection lines in survey KR13-11 along the Japan trench. **A.** Sediment accretion is promoted in regions where sediment thickness on the incoming plate is large relative to incoming plate fault throw on outer-rise faults (ratio >1). In this scenario, the décollement undulates with subducting horsts and grabens and may remain in the basal clay layer. **B.** Frontal tectonic erosion is promoted in regions with small sediment thickness relative to fault throw (ratio <1). At the Japan trench, the presence of petit spot volcanism may contribute slope failures at the deformation front and facilitate frontal tectonic erosion.

efficients than the overlying biogenic mudstones (Ikari et al., 2015). Correlation of cores to seismic reflection lines imply that this clay layer is present across the study area, near the base of seismic unit SU2 (Nakamura et al., 2013) (Figure 1, Core C0019), except in regions where there is petit spot volcanism (Fujie et al., 2020). Therefore, the presence of this frictionally weak clay may help promote the development of undulating décollements that step into grabens in regions of low fault throw, but segmentation of the clay layer may hinder this process, leading to a planar décollement and compositional décollement heterogeneity.

Our data therefore provide important insights into the spatial scales of décollement heterogeneity that occur in the shallow subduction interface. Our frontal prism mapping suggests that relatively compositionally homogenous patches of décollement hosted in frictionally weak clays may occur in 5–20 km wide by 15–40 km long regions experiencing sediment accretion (Figure 6d). These homogeneous regions are likely segmented by ~5 km wide by 5–30 km long patches, or potential asperities, where the décollement is hosted in frictionally stronger biogenic muds and frontal prism material, in locations where partial accretion, sediment subduction, and tectonic erosion are mapped (Figure 6d).

Similar patterns can be observed in the interpolated map of incoming plate ratios (Figure 6d). This map shows 5–20 km wide by 20 to >40 km long laterally continuous regions of crust with ratios >1 that will likely promote sediment accretion and the formation of compositionally homogenous décollement patches. These patches are segmented by regions with ratios <1, where deformation mode, and therefore décollement mechanics, will likely vary over ~5–10 km length scales. These length scales are similar to those observed on the interpolated map of frontal prism deformation mode (Figure 6d). The incoming plate ratio map, therefore, may serve as a tool for interpreting patterns and length scales of spatio-temporal heterogeneity on the shallow plate interface caused by variations in décollement geometry and composition.

5.3 Implications for slip potential

The compositional and frictional properties of the décollement in northeast Japan have been proposed to be important factors in accommodating shallow seismogenic slip to the trench (e.g., Chester et al., 2013; Kameda et al., 2015; Moore et al., 2015). In particular, structural and lithological descriptions of cores and borehole logs crossing the plate boundary (Chester et al., 2013), interpretations of high-resolution seismic reflection data across the trench (Nakamura et al., 2013), and frictional heating across the plate boundary (Fulton et al., 2013) suggest that the Tohoku earthquake slip surface localized in the basal zone of frictionally weak, smectite-rich pelagic clay. Additionally, published seismic reflection data collected across the portion of the earthquake with the greatest amount of slip show a deformation front with imbricate thrust faults that sole into the décollement, positioned near the base of the sediment section, that undulates with subducting horsts and grabens (Boston et al., 2014; Chester and Moore, 2018; Nakamura et al., 2013, 2020). The décollement and the branching imbricate thrust faults have been suggested to be the shallowest faults that hosted Tohoku earthquake slip (Nakamura et al., 2020).

The degree of heterogeneity along the plate interface imparted by outer-rise normal fault subduction may influence seismogenic slip to the trench. Specifically, we find that in the southern portion of the map area (south of 39°N) where shallow slip occurred during the 2011 Tohoku earthquake, incoming plate sediment thickness is mainly greater than fault throw (ratio >1) and complete or partial sediment accretion are the predominant modes of deformation. These data suggest that the upper ~20 km of the shallow subduction interface in this region may have a relatively lithologically homogenous and frictionally weak composition that promotes shallow seismogenic slip. These interpretations agree with prior work which demonstrates that the region of large slip during the Tohoku earthquake occurred in regions experiencing sediment accretion via imbricate thrust faulting (Nakamura et al., 2020).

Conversely, in the northern portion of the map area (~39.1–39.7°N), there is greater variability in the mapped deformation modes and lower ratios of sediment thickness to fault throw at the trench. These data suggest the décollement here may be compositionally heterogeneous and may be developed in frictionally stronger materials. This portion of the margin is known to host tectonic tremors and transient aseismic slip (e.g., Nishikawa et al., 2023). We identify two possible sources of heterogeneity in the shallow plate interface in this region. First, the incoming plate outboard of this region contains thin sediments and petit spot volcanism (Figure 5). Petit spot volcanism is thought to thermally metamorphose the incoming plate sediment section, which alters the compositional properties and may increase the friction of these biogenic muds and clays (Fujie et al., 2020). Therefore, in these locations, the compositional and frictional properties of the incoming plate sediments are different than the surrounding, unaltered incoming plate sediments, and may disrupt décollement development and limit slip potential in these areas. Second, this region contains large-offset normal faults on the incoming plate (ratios <1) that may also disrupt décollement development by introducing large geometric asperities and by promoting lateral heterogeneity in the composition and frictional properties of the sediments in which the décollement develops. Such geometric asperities and lateral variations in compositional and frictional properties of the décollement have been thought to inhibit large magnitude slip to the trench (Kodaira et al., 2019; Fujie et al., 2020; Moore et al., 2015; Qin et al., 2022). Therefore, our results demonstrate that sediment thickness and outer-rise fault throw may exert a significant control on the composition, friction, and geometric heterogeneity of the décollement that may help promote or inhibit large magnitude, shallow tsunamigenic slip. These relationships have important implications for slip potential at the Japan trench as well as in sediment-starved subduction systems globally where outer-rise normal faults are subducted.

6 Conclusion

Outer-rise faults and sediment thickness on the incoming plate are direct inputs into the shallow subduction zone, and therefore have an important influence on frontal prism modes of deformation, décollement evolution, and the potential for shallow plate boundary slip. We mapped sediment thickness on the incoming plate, the amount of fault throw across normal faults that bound horsts and grabens, instances of petit spot volcanism on the incoming plate, and slope failures at the deformation front for a portion of the Japan trench. We show heterogeneity in the modes of frontal prism deformation at 5–10 km length scales along strike and down dip. We find that portions of the incoming plate where sediment thickness is greater than fault throw may promote sediment accretion and the development of an undulating, lithologically homogenous décollement. Conversely, subduction of incoming plate segments with a thin sediment section and high offset faults may promote tectonic erosion and slope failures, and the development of a planar but lithologically heterogeneous décollement that smooths over subducting horsts and grabens.

The degree of heterogeneity observed in the frontal prism deformation mode requires correlative geometric and compositional heterogeneity in the shallow décollement and may have important implications for the mechanics of and potential for shallow plate boundary slip. In particular, the 2011 Tohoku earthquake ruptured the southern portion of the study area, where outer-rise fault throws are small relative to sediment thickness and sediment accretion is the dominant mode of deformation and ended in a region that transitions to high fault throw and thin sediments. Because the Japan trench has a range of sediment thicknesses and fault throws that are characteristic to many sediment-starved margins around the globe, we propose that the ratio of sediment thickness to fault throw may be a useful proxy for understanding how outer-rise fault throw and incoming sediment thickness impact frontal prism deformation style, sediment flux, and décollement heterogeneity at other margins.

Acknowledgements

Funding for this project was provided by the National Science Foundation grant OCE-2103514 and the Northern Arizona University Duebendorfer & Barnes Structural Geology Endowment and Eminent Scholars Foundations. Data collection and processing were supported by the Japan Agency for Marine-Earth Science and Technology (JASMTEC). We thank IHS S&P Global for providing academic licenses for their Kingdom Suite software. We also thank Rebecca Bell and Harold Tobin for their constructive reviews that helped improve this manuscript.

7 Data and code availability

The seismic reflections profiles used in this study can be requested and accessed from https://www.jamstec.go.jp/obsmcs_db/e/. Data tables in this study can be found in the Supporting Information and on Zenodo (Schottenfels et al., 2023).

8 Competing interests

The authors have no competing interests.

References

Boston, B., Moore, G. F., Nakamura, Y., and Kodaira, S. Outer-rise normal fault development and influence on near-trench décollement propagation along the Japan Trench, off Tohoku. *Earth, Planets and Space*, 66(1):135, 2014. doi: 10.1186/1880-5981-66-135.

Chester, F. M. and Moore, J. C. *Tectonostratigraphy and processes of frontal accretion with horst-graben subduction at the Japan Trench*. Geological Society of America, 2018. doi: 10.1130/2018.2534(06).

Chester, F. M., Rowe, C., Ujiie, K., Kirkpatrick, J., Regalla, C., Remitti, F., Moore, J. C., Toy, V., Wolfson-Schwehr, M., Bose, S., Kameda, J., Mori, J. J., Brodsky, E. E., Eguchi, N., and Toczko, S. Structure and Composition of the Plate-Boundary Slip Zone for the 2011 Tohoku-Oki Earthquake. *Science*, 342(6163):1208–1211, Dec. 2013. doi: 10.1126/science.1243719.

Clift, P. and Vannucchi, P. Controls on tectonic accretion versus erosion in subduction zones: Implications for the origin and recycling of the continental crust. *Reviews of Geophysics*, 42(2), Apr. 2004. doi: 10.1029/2003rg000127.

Contreras-Reyes, E., Grevemeyer, I., Flueh, E. R., Scherwath, M., and Heesemann, M. Alteration of the subducting oceanic lithosphere at the southern central Chile trench–outer rise. *Geochemistry, Geophysics, Geosystems*, 8(7), July 2007. doi: 10.1029/2007gc001632.

Fujie, G., Kodaira, S., Nakamura, Y., Morgan, J. P., Dannowski, A., Thorwart, M., Grevemeyer, I., and Miura, S. Spatial variations of incoming sediments at the northeastern Japan arc and their implications for megathrust earthquakes. *Geology*, 48(6):614–619, Mar. 2020. doi: 10.1130/g46757.1.

Fujiwara, T., Hirano, N., Abe, N., and Takizawa, K. Subsurface structure of the "petit-spot" volcanoes on the northwestern Pacific Plate. *Geophysical Research Letters*, 34(13), July 2007. doi: 10.1029/2007gl030439.

Fulton, P. M., Brodsky, E. E., Kano, Y., Mori, J., Chester, F., Ishikawa, T., Harris, R. N., Lin, W., Eguchi, N., and Toczko, S. Low Coseismic Friction on the Tohoku-Oki Fault Determined from Temperature Measurements. *Science*, 342(6163):1214–1217, Dec. 2013. doi: 10.1126/science.1243641.

Hilde, T. W. Sediment subduction versus accretion around the pacific. *Tectonophysics*, 99(2–4):381–397, Dec. 1983. doi: 10.1016/0040-1951(83)90114-2.

Hirano, N., Takahashi, E., Yamamoto, J., Abe, N., Ingle, S. P., Kaneoka, I., Hirata, T., Kimura, J.-I., Ishii, T., Ogawa, Y., Machida, S., and Suyehiro, K. Volcanism in Response to Plate Flexure. *Science*, 313(5792):1426–1428, Sept. 2006. doi: 10.1126/science.1128235.

Hirano, N., Machida, S., Sumino, H., Shimizu, K., Tamura, A., Morishita, T., Iwano, H., Sakata, S., Ishii, T., Arai, S., Yoneda, S., Danhara, T., and Hirata, T. Petit-spot volcanoes on the oldest portion of the Pacific plate. *Deep Sea Research Part I: Oceanographic Research Papers*, 154:103142, Dec. 2019. doi: 10.1016/j.dsr.2019.103142.

Iinuma, T., Hino, R., Kido, M., Inazu, D., Osada, Y., Ito, Y., Ohzono, M., Tsushima, H., Suzuki, S., Fujimoto, H., and Miura, S. Coseismic slip distribution of the 2011 off the Pacific Coast of Tohoku Earthquake (M9.0) refined by means of seafloor geodetic data. *Journal of Geophysical Research: Solid Earth*, 117(B7), July 2012. doi: 10.1029/2012jb009186.

Ikari, M. J., Kameda, J., Saffer, D. M., and Kopf, A. J. Strength characteristics of Japan Trench borehole samples in the high-slip region of the 2011 Tohoku-Oki earthquake. *Earth and Planetary Science Letters*, 412:35–41, Feb. 2015. doi: 10.1016/j.epsl.2014.12.014.

Ikehara, K., Usami, K., Kanamatsu, T., Arai, K., Yamaguchi, A., and Fukuchi, R. Spatial variability in sediment lithology and sedimentary processes along the Japan Trench: use of deep-sea turbidite records to reconstruct past large earthquakes. *Geological Society, London, Special Publications*, 456(1):75–89, Mar. 2017. doi: 10.1144/sp456.9.

Kameda, J., Shimizu, M., Ujiie, K., Hirose, T., Ikari, M., Mori, J., Oohashi, K., and Kimura, G. Pelagic smectite as an important factor in tsunamigenic slip along the Japan Trench. *Geology*, 43 (2):155–158, Feb. 2015. doi: 10.1130/g35948.1.

Kirkpatrick, J. D., Rowe, C. D., Ujiie, K., Moore, J. C., Regalla, C., Remitti, F., Toy, V., Wolfson-Schwehr, M., Kameda, J., Bose, S., and Chester, F. M. Structure and lithology of the Japan Trench subduction plate boundary fault. *Tectonics*, 34(1):53–69, Jan. 2015. doi: 10.1002/2014tc003695.

Kodaira, S., No, T., Nakamura, Y., Fujiwara, T., Kaiho, Y., Miura, S., Takahashi, N., Kaneda, Y., and Taira, A. Coseismic fault rupture at the trench axis during the 2011 Tohoku-oki earthquake. *Nature Geoscience*, 5(9):646–650, Aug. 2012. doi: 10.1038/ngeo1547.

Kodaira, S., Nakamura, Y., Yamamoto, Y., Obana, K., Fujie, G., No, T., Kaiho, Y., Sato, T., and Miura, S. Depth-varying structural characters in the rupture zone of the 2011 Tohoku-oki earthquake. *Geosphere*, 13(5):1408–1424, Sept. 2017. doi: 10.1130/ges01489.1.

Kodaira, S., Nakamura, Y., Fujie, G., and Miura, S. Marine active-source seismic studies in the Japan Trench: a seismogenic zone in an ocean-continent collision zone. *Acta Geologica Sinica - English Edition*, 93(S1):94–95, May 2019. doi: 10.1111/1755-6724.13959.

Kodaira, S., Fujiwara, T., Fujie, G., Nakamura, Y., and Kanamatsu, T. Large Coseismic Slip to the Trench During the 2011 Tohoku-Oki Earthquake. *Annual Review of Earth and Planetary Sciences*, 48(1):321–343, May 2020. doi: 10.1146/annurev-earth-071719-055216.

Masson, D. G. Fault patterns at outer trench walls. *Marine Geophysical Researches*, 13(3):209–225, Aug. 1991. doi: 10.1007/bf00369150.

Moore, G. F., Shipley, T. H., and Lonsdale, P. F. Subduction erosion versus sediment offscraping at the toe of the Middle America Trench off Guatemala. *Tectonics*, 5(4):513–523, Aug. 1986. doi: 10.1029/tc005i004p00513.

Moore, J. C., Plank, T. A., Chester, F. M., Polissar, P. J., and Savage, H. M. Sediment provenance and controls on slip propagation: Lessons learned from the 2011 Tohoku and other great earthquakes of the subducting northwest Pacific plate. *Geosphere*, 11(3):533–541, Apr. 2015. doi: 10.1130/ges01099.1.

Morgan, J. K., Ramsey, E. B., and Ask, M. V. S. *8. Deformation and Mechanical Strength of Sediments at the Nankai Subduction Zone: Implications for Prism Evolution and Décollement Initiation and Propagation*, page 210–256. Columbia University Press, Dec. 2007. doi: 10.7312/dixo13866-008.

Nakamura, Y., Kodaira, S., Miura, S., Regalla, C., and Takahashi, N. High-resolution seismic imaging in the Japan Trench axis area off Miyagi, northeastern Japan. *Geophysical Research Letters*, 40 (9):1713–1718, May 2013. doi: 10.1002/grl.50364.

Nakamura, Y., Fujiwara, T., Kodaira, S., Miura, S., and Obana, K. Correlation of frontal prism structures and slope failures near the trench axis with shallow megathrust slip at the Japan Trench. *Scientific Reports*, 10(1), July 2020. doi: 10.1038/s41598-020-68449-6.

Nakamura, Y., Kodaira, S., Fujie, G., Yamashita, M., Obana, K., and Miura, S. Incoming plate structure at the Japan Trench subduction zone revealed in densely spaced reflection seismic profiles. *Progress in Earth and Planetary Science*, 10(1), Aug. 2023. doi: 10.1186/s40645-023-00579-7.

Nasu, N., von Huene, R., Ishiwada, Y., Langseth, L., Bruns, T., and Honza, E. *Interpretation of Multichannel Seismic Reflection Data, Legs 56 and 57, Japan Trench Transect, Deep Sea Drilling Project*. U.S. Government Printing Office, Nov. 1980. doi: 10.2973/dsdp.proc.5657.112.1980.

Nishikawa, T., Ide, S., and Nishimura, T. A review on slow earthquakes in the Japan Trench. *Progress in Earth and Planetary Science*, 10(1), Jan. 2023. doi: 10.1186/s40645-022-00528-w.

Polet, J. and Kanamori, H. Shallow subduction zone earthquakes and their tsunamigenic potential. *Geophysical Journal International*, 142(3):684–702, Sept. 2000. doi: 10.1046/j.1365-246x.2000.00205.x.

Qin, Y., Nakamura, Y., Kodaira, S., and Fujie, G. Seismic imaging of subsurface structural variations along the Japan trench south of the 2011 Tohoku earthquake rupture zone. *Earth and Planetary Science Letters*, 594:117707, Sept. 2022. doi: 10.1016/j.epsl.2022.117707.

Regalla, C., Bierman, P., and Rood, D. H. Meteoric 10Be Reveals a

Young, Active Accretionary Prism and Structurally Complex Décollement in the Vicinity of the 2011 Tohoku Earthquake Rupture. *Geochemistry, Geophysics, Geosystems*, 20(11):4956–4971, Nov. 2019. doi: 10.1029/2019gc008483.

Saffer, D. M. and Tobin, H. J. Hydrogeology and Mechanics of Subduction Zone Forearcs: Fluid Flow and Pore Pressure. *Annual Review of Earth and Planetary Sciences*, 39(1):157–186, May 2011. doi: 10.1146/annurev-earth-040610-133408.

Sawai, M., Niemeijer, A. R., Hirose, T., and Spiers, C. J. Frictional properties of JFAST core samples and implications for slow earthquakes at the Tohoku subduction zone. *Geophysical Research Letters*, 44(17):8822–8831, Sept. 2017. doi: 10.1002/2017gl073460.

Schottenfels, E., Regalla, C., and Nakamura, Y. Influence of outer-rise faults on sediment flux and décollement heterogeneity at the Japan trench, 2023. doi: 10.5281/zenodo.7683455.

Seno, T., Stein, S., and Gripp, A. E. A model for the motion of the Philippine Sea Plate consistent with NUVEL-1 and geological data. *Journal of Geophysical Research: Solid Earth*, 98(B10): 17941–17948, Oct. 1993. doi: 10.1029/93jb00782.

Shipboard Scientific Party. *Site 436, Japan Trench Outer Rise, Leg 56*. U.S. Government Printing Office, Nov. 1980. doi: 10.2973/dsdp.proc.5657.107.1980.

Strasser, M., Kölling, M., Ferreira, C. d. S., Fink, H., Fujiwara, T., Henkel, S., Ikehara, K., Kanamatsu, T., Kawamura, K., Kodaira, S., Römer, M., and Wefer, G. A slump in the trench: Tracking the impact of the 2011 Tohoku-Oki earthquake. *Geology*, 41(8): 935–938, Aug. 2013. doi: 10.1130/g34477.1.

Sun, T., Saffer, D., and Ellis, S. Mechanical and hydrological effects of seamount subduction on megathrust stress and slip. *Nature Geoscience*, 13(3):249–255, Mar. 2020. doi: 10.1038/s41561-020-0542-0.

Tanioka, Y., Ruff, L., and Satake, K. What controls the lateral variation of large earthquake occurrence along the Japan Trench? *Island Arc*, 6(3):261–266, Sept. 1997. doi: 10.1111/j.1440-1738.1997.tb00176.x.

Tsuru, T., Park, J., Takahashi, N., Kodaira, S., Kido, Y., Kaneda, Y., and Kono, Y. Tectonic features of the Japan Trench convergent margin off Sanriku, northeastern Japan, revealed by multichannel seismic reflection data. *Journal of Geophysical Research: Solid Earth*, 105(B7):16403–16413, July 2000. doi: 10.1029/2000jb900132.

Tsuru, T., Park, J., Miura, S., Kodaira, S., Kido, Y., and Hayashi, T. Along-arc structural variation of the plate boundary at the Japan Trench margin: Implication of interplate coupling. *Journal of Geophysical Research: Solid Earth*, 107(B12), Dec. 2002. doi: 10.1029/2001jb001664.

von Huene, R. and Culotta, R. Tectonic erosion at the front of the Japan Trench convergent margin. *Tectonophysics*, 160(1–4): 75–90, Mar. 1989. doi: 10.1016/0040-1951(89)90385-5.

von Huene, R. and Lallemand, S. Tectonic erosion along the Japan and Peru convergent margins. *Geological Society of America Bulletin*, 102(6):704–720, June 1990. doi: 10.1130/0016-7606(1990)102<0704:teatja>2.3.co;2.

von Huene, R., Langseth, M., Nasu, N., and Okada, H. A summary of Cenozoic tectonic history along the IPOD Japan Trench transect. *Geological Society of America Bulletin*, 93(9):829, 1982. doi: 10.1130/0016-7606(1982)93<829:asocth>2.0.co;2.

Wang, K. and Bilek, S. L. Invited review paper: Fault creep caused by subduction of rough seafloor relief. *Tectonophysics*, 610: 1–24, Jan. 2014. doi: 10.1016/j.tecto.2013.11.024.

doi:10.26443/seismica.v3i1.367

SEISMICA

Ocean bottom seismometer clock correction using ambient seismic noise

David Naranjo * [1,2], **Laura Parisi** [2], **Sigurjón Jónsson** [2], **Philippe Jousset** [3], **Dieter Werthmüller** [1], **Cornelis Weemstra** [1,4]

[1]Department of Geoscience and Engineering, Delft University of Technology, Stevinweg 1, 2628 CN, Delft, the Netherlands, [2]Physical Science and Engineering Division, King Abdullah University of Science and Technology, Thuwal, Makkah 23955, Saudi Arabia, [3]GFZ German Research Centre for Geosciences, Helmholtz Centre Potsdam, Section 2.2 Geophysical Imaging, Potsdam, Germany, [4]Department of Seismology and Acoustics, Royal Netherlands Meteorological Institute, Utrechtseweg 297, 3730 AE, De Bilt, the Netherlands

Author contributions: *Conceptualization*: CW, DN. *Software*: DN, CW, DW. *Formal Analysis*: All authors. *Investigation*: All authors. *Writing - original draft*: DN. *Writing - Review & Editing*: All authors. *Data acquisition*: PJ.

Abstract Ocean-bottom seismometers (OBSs) are equipped with seismic sensors that record acoustic and seismic events at the seafloor. One critical parameter for obtaining accurate earthquake locations is the absolute time of the recorded seismic signals. It is, however, not possible to synchronize the internal clocks of the OBSs with a known reference time, as GNSS signals do not reach the sea bottom. We address this issue by introducing a new method to synchronize the clocks of large-scale OBS deployments. Similar to some previous approaches, our method leverages the theoretical time-symmetry of time-averaged cross-correlations of ambient seismic noise: broken time-symmetry is attributed to clock drift. A non-uniform surface wave illumination pattern, however, can also break the time-symmetry. Existing noise-based synchronization techniques usually ignore the latter, but we do address it by means of a weighted least-squares inversion (based on station-to-station distances). The weighted least-squares inversion mitigates the adverse effect of a non-uniform surface wave illumination on the time-symmetry. Furthermore, our method includes a unique feature: it estimates and corrects for an initial clock error introduced at the deployment time. This initial clock error can be attributed to either (i) a wrong initial time synchronization or (ii) the temperature shock during deployment. The methodology is implemented in an open-source Python package named `OCloC` and was tested with OBS recordings acquired around the Reykjanes peninsula, southwest Iceland. Our results indicate that all OBSs experienced a clock drift, and that a significant number of them were subject to an initial clock error at the deployment time. This study provides a substantial improvement in the inherent quality of OBS data, laying a solid foundation for more robust seismic data analysis.

Production Editor:
Gareth Funning
Handling Editor:
Lise Retailleau
Copy & Layout Editor:
Abhineet Gupta

Received:
February 5, 2023
Accepted:
November 29, 2023
Published:
January 19, 2024

Non-technical summary Ocean-bottom seismometers (OBSs) are instruments deployed on the seafloor, equipped with sensors to record seismic activity offshore. However, getting accurate information from these instruments is challenging because the internal clocks of the OBSs cannot be easily synchronized with a known reference time. In this study, we developed a new approach to synchronize the clocks of large-scale OBS deployments. Our approach uses cross-correlations of ambient seismic noise to detect errors in the timing of the sensor clocks. We implemented our methodology in the open-source Python package `OCloC` and tested it on data from a seismic network deployed offshore the Reykjanes peninsula, southwest Iceland. This new approach will aid in improving the accuracy of earthquake locations and imaging the crust and upper mantle.

1 Introduction

Over the past few decades, there has been an increase in the use of ocean-bottom seismometers (OBSs). OBS readings allow one to identify remarkable features such as undersea volcanic eruptions (Matsumoto et al., 2019) or seismic activity linked to tectonic strain and gas emissions through fault conduits (Tary et al., 2011). In particular, OBS readings are frequently used for imaging of the crust and/or mantle (e.g., Dongmo Wamba et al., 2023). Despite these successes, a key challenge in using OBSs remains the accurate (time) synchronization of the instruments' recordings. In fact, most OBS clocks drift, meaning they do not run at the same rate as a reference clock. This issue might be overcome by using atomic clocks instead of the traditional microprocessor-compensated crystal oscillator clocks that most OBSs have (Gardner and Collins, 2012). This, however, would increase the inventory costs and power consumption, implying fewer instruments and less monitoring time,

*Corresponding author: d.f.naranjohernandez@tudelft.nl

respectively. If the network is not properly synchronized, the incorrectly timed recordings may result in biased earthquake locations and Earth structure models.

One simple approach to identify clock drift is to measure the time difference between the instrument's internal clock and a GNSS signal before deployment and after recovery. This time difference is commonly referred to as the instrument's 'skew'. Assuming the instrument's clock drifted at a linear rate, a time correction can then be applied (e.g., Geissler et al., 2010). The skew, however, is not always possible to retrieve (e.g., when the instrument's battery dies before recovery). For this reason, several authors have proposed alternative methods for correcting clock errors; many of these exploiting the presumed temporal stability of time-averaged cross-correlations of ambient seismic noise (e.g., Sens-Schönfelder, 2008; Loviknes et al., 2020; Hannemann et al., 2014; Jousset et al., 2013). These approaches, however, ignore errors that could arise if the initial synchronization with a GNSS signal is either lacking or erroneous, or if there is an "initial" clock error resulting from the temperature shock during deployment (Zhang et al., 2023).

In theory, time-averaged cross-correlations of recordings of ambient seismic noise (henceforth 'noise cross-correlations') result in a signal that is symmetric around $t = 0$ (e.g., Stehly et al., 2006). In fact, under favorable conditions, the signals at positive and negative time lag coincide with the medium's Green's function (between the positions of the two seismic stations) and its time reverse, respectively. As such, it is referred to as 'seismic interferometry' (SI) (Wapenaar and Fokkema, 2006). In practice, these conditions are often not entirely fulfilled. Notwithstanding, provided the illumination is sufficiently uniform, the operation of averaging noise cross-correlations over time still yields two interferometric surface wave responses: one at the positive and one at the negative time lag(s). Violation of the noise cross-correlations' time symmetry may indicate the presence of clock errors (e.g., Hannemann et al., 2014).

Currently, two distinct approaches use noise cross-correlations to detect and correct clock errors (Gouédard et al., 2014). The first approach is based on the presumed temporal stability of the noise cross-correlations (Hable et al., 2018; Loviknes et al., 2020). In this approach, cross-correlation functions (CCFs) of ambient noise are calculated over different periods. The drift is then estimated as the time shift that maximizes the Pearson correlation coefficient between each CCF and a reference correlation function (Hable et al., 2018). However, this method ignores the possibility of an initial clock error at the time of deployment due to a temperature shock during the OBS' descent to the ocean floor (Gardner and Collins, 2012; Zhang et al., 2023). The second approach exploits the above-mentioned time symmetry between the retrieved interferometric responses (Sens-Schönfelder, 2008; Weemstra et al., 2021). Contrary to the first approach, both direct surface wave arrivals (i.e., at positive and negative time lag(s) need to be retrieved successfully in this case. Low signal-to-noise ratios or stations that are too close to each other (in terms of wavelength) prohibit this.

Although existing approaches for correcting clock errors have proven successful, a few challenges remain. First, the symmetry of ambient noise cross-correlation, while a valuable theoretical concept, is rarely realized in practice. A non-uniform illumination pattern may cause shifts in the arrival time of the interferometric responses with respect to the true arrival time (a challenge that is often overlooked). Second, current methods ignore the possibility of the aforementioned initial clock error during deployment. This clock error, introduced during the OBS' descent, is not expected given the mechanism causing clock drift (e.g., Shariat-Panahi et al., 2009), but it would nonetheless be good to rule out; in particular because the first approach mentioned above (Hable et al., 2018; Loviknes et al., 2020), does not allow such initial clock error to be detected. Finally, many of the current methods rely on land seismometers that are considered to be devoid of clock errors, ideally in the vicinity of the OBS deployment. This, however, will not be the case when the OBS network is located in oceanic regions far from the coast.

In this paper, we present a versatile method that addresses all these challenges. Our approach (i) uses a weighted least-squares inversion to minimize the detrimental effect of non-uniform illumination patterns, (ii) allows for a potential initial clock error at deployment time, and (iii) does not require land stations to be included in the network to synchronize the recordings. Regarding the third claim, although our approach allows the OBS network's recordings to be synchronized, the combined set of recordings cannot be synchronized with Coordinated Universal Time (UTC). To achieve that, a land station (with a UTC-synchronized clock) needs to be included in the network. The presented method is implemented in an open-source Python package named `OCloC` (OBS Clock Correction), which accompanies this paper. It combines the two aforementioned techniques for clock error detection (i.e., the one relying on the presumed temporal stability of noise cross-correlations and the one relying on their presumed time symmetry). Our method (and hence the package) is particularly useful in application to large-N seismic arrays.

To show the validity of our method, we use data from a seismic network deployed on and around the Reykjanes peninsula, SW Iceland (Jousset et al., 2020a). This seismic network was deployed in the context of the geothermal project IMAGE (Integrated Methods for Advanced Geothermal Exploration, see also Jousset et al., 2020b; Blanck et al., 2020). The data set used consists of recordings by 30 on-land stations and 17 OBSs (this is a subset of the stations used in Weemstra et al., 2021). In the following sections, we detail the theory underlying our approach (Section 2), discuss and exemplify the implementation of this theory (Section 3), present and discuss our findings (Section 4), and list the most important conclusions (Section 6). A more detailed description of the data is included in Section 3 (Section 3.1). In addition, a brief description of `OCloC` is given in this section (Section 3.3).

2 Theory

In this section, the theory is introduced step-wise. First, we briefly highlight the most important theoretical aspects of Seismic Interferometry (SI). Second, we introduce a model adequate for determining clock drift, which is an extension of the model introduced by Weemstra et al. (2021). Third, we introduce potential additional time shifts (i.e., in addition to clock drift) affecting the arrival times of the interferometric responses. Fourth, we describe how a single noise cross-correlation's drift, and deviation from symmetry, can be retrieved. Fifth, we present the matrix notation of the introduced model. Finally, we briefly describe two different inversion approaches.

2.1 Seismic interferometry

Early types of seismic interferometry (SI) were introduced to the geophysics community by Aki (1957) and Claerbout (1968). Over the last two decades, the theory underlying SI has been established (Lobkis and Weaver, 2001; Wapenaar and Fokkema, 2006; Snieder, 2004; Shapiro and Campillo, 2004), and the method has been exploited in numerous applications. Examples include subsurface characterization (Draganov et al., 2007; Jousset et al., 2016), reservoir monitoring (Sánchez-Pastor et al., 2019), and glaciology (Lindner et al., 2018). In this study, SI is used as an independent method to recover clock errors without needing skew measurements.

Applying SI to recordings of ambient seismic noise allows one to retrieve new seismic responses between pairs of stations by means of simple cross-correlations (Wapenaar and Fokkema, 2006; Stehly et al., 2006). Under specific conditions, the time-averaged cross-correlation contains the response to two 'virtual sources': one at negative lag times (usually referred to as the 'acausal part') and another at positive lag times (referred to as the 'causal part'), and with the virtual sources coinciding with the receiver locations. Time averaging is required to suppress spurious travel time delays that arise from constructive interference of signals coming from different sources. The time-averaged noise cross-correlation is proportional to the medium's Green's function if: (i) the noise sources illuminate the station pairs uniformly from all angles, (ii) the noise sources are uncorrelated, (iii) the medium is lossless, and (iv) sources have coinciding amplitude spectra (Wapenaar and Fokkema, 2006). Under these assumptions, the time-averaged cross-correlation of noise recorded by stations at $\mathbf{x}_i$ and $\mathbf{x}_j$, which we denote by $C_{i,j}(t)$, is proportional to the Green's function $G(\mathbf{x}_j, \mathbf{x}_i, t)$ and its time-reversed version, convolved with the autocorrelation of the signal emitted by the (noise) sources, i.e.,

$$C_{i,j}(t) \propto [G(\mathbf{x}_j, \mathbf{x}_i, t) + G(\mathbf{x}_j, \mathbf{x}_i, -t)] * P(t), \quad (1)$$

where $P(t)$ denotes the signal's autocorrelation generated by noise sources. In this study, we focus on the direct surface wave part of the Green's functions, ignoring the scattered signal. We refer to Wapenaar and Fokkema (2006) and Halliday and Curtis (2008) for a more detailed discussion of the assumptions underlying SI.

2.2 A model to account for clock drift

When it comes to the recovery of clock errors, an essential feature of the noise cross-correlation is its presumed time symmetry: under the assumptions listed in the previous section, the direct surface waves in $C_{i,j}(t)$ arrive at time lags of equal magnitude but opposite signs (Figure 1a). A violation of this time symmetry, such as the one in Figure 1b, indicates the presence of clock errors. To infer these clock errors from noise cross-correlations, Weemstra et al. (2021, Section 4) recently introduced an appropriate model. These authors, however, did not include clock drift in their model as they assumed the instrumental clock errors to be time-independent (or constant). We extend the model introduced by Weemstra et al. (2021) to account for time-dependent clock errors such as clock drift.

Here we assume the (potential) OBS clock drift to be linear. This is based on the fact that (i) the drift rate should be steady at constant temperature and (ii) the ambient temperature tends to be rather stable in deep water (note that the drift rate at a certain temperature is dictated by the frequency of the quartz oscillators in seismic clocks; Shariat-Panahi et al., 2009). The validity of this assumption has been demonstrated for OBSs at larger depths in previous studies (Hable et al., 2018; Loviknes et al., 2020).

To estimate clock drift, we compute time-lapse cross-correlations $C_{i,j}(t, t^{(\mathrm{lps})})$, where $t^{(\mathrm{lps})}$ is the timing of the time-lapse cross-correlation. We refer to $C_{i,j}(t, t^{(\mathrm{lps})})$ as the 'lapse cross-correlation'. Note that $t^{(\mathrm{lps})}$ is the average time of all time windows contributing to the lapse cross-correlation. Therefore, $t^{(\mathrm{lps})}$ is not necessarily the time exactly in between the time of the first and last time window contributing to $C_{i,j}(t, t^{(\mathrm{lps})})$: in case the recordings by one of the two stations (or both) contain gaps, $t^{(\mathrm{lps})}$ may be skewed towards the beginning or end of the entire period over which individual cross-correlations are averaged.

For the considered linear parametrization, the time-dependent clock error of station i, denoted by $\delta t_i^{(\mathrm{ins})}$, is written as

$$\delta t_i^{(\mathrm{ins})}\left(t^{(\mathrm{lps})}\right) = a_i t^{(\mathrm{lps})} + b_i, \quad (2)$$

where $\delta t_i^{(\mathrm{ins})}$ is the clock error of station i at $t^{(\mathrm{lps})}$, $t^{(\mathrm{lps})}$ is the average time of the time-lapse cross-correlation, a_i is the clock drift rate of station i, and b_i is the incurred clock error of station i at $t^{(\mathrm{lps})} = 0$.

Note that $t^{(\mathrm{lps})}$ is a continuous variable and that it is conveniently (but arbitrarily) set to 0 at the 21st of August 2014. This is the approximate time of deployment of the OBSs considered in this study (the OBSs have been deployed over the course of a number of days around that date). Furthermore, $\delta t_i^{(\mathrm{ins})}$ is defined such that negative values imply that the recordings by station i are subject to a time delay. The rate at which the clock of station i is drifting is given by a_i, whereas b_i represents

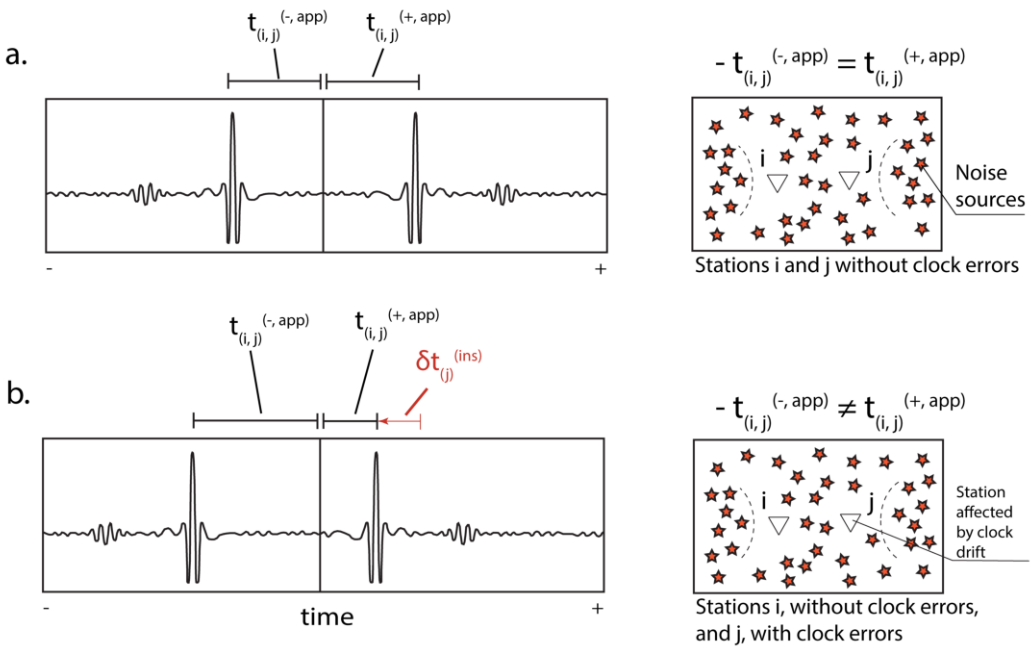


Figure 1 a. Noise cross-correlations computed using two stations without clock errors. The noise cross-correlation is almost symmetric in this case (for a relatively uniform illumination), and $t_{i,j}^{(+,\text{app})} = -t_{i,j}^{(-,\text{app})}$. b. Noise cross-correlations computed while one of the two stations is subject to clock errors (e.g., due to clock drift of one or both instruments). The noise cross-correlation is asymmetric (even for a relatively uniform illumination), and $t_{i,j}^{(+,\text{app})} \neq -t_{i,j}^{(-,\text{app})}$. In b, station j is subject to a clock error of $\delta t_j^{(\text{ins})}$, which causes the noise cross-correlation to shift to negative time by that amount.

a possible clock error of station i at $t^{(\text{lps})} = 0$. These are the two unknown parameters that we want to recover in this study (for all the OBSs). A different parametrization of $\delta t_i^{(\text{ins})}$ in terms of, for example, cubic splines or trigonometric basis functions (i.e., Fourier series) is relatively straightforward.

A deviation from time symmetry can result from clock errors in either one or both stations involved in the noise cross-correlation. Similar to Weemstra et al. (2021), we denote the arrival time of the causal direct surface wave in $C_{i,j}(t, t^{(\text{lps})})$ by $t_{i,j}^{(+,\text{app})}$ and the arrival time of the acausal direct surface wave by $t_{i,j}^{(-,\text{app})}$. Accounting for the time-dependent clock errors above, we obtain the following expression for the apparent arrival time of the causal direct surface wave:

$$t_{i,j}^{(+,\text{app})}\left(t^{(\text{lps})}\right) = t_{i,j}^{(+)} + \delta t_i^{(\text{ins})}\left(t^{(\text{lps})}\right) - \delta t_j^{(\text{ins})}\left(t^{(\text{lps})}\right). \tag{3}$$

Similarly, the apparent arrival time of the acausal direct surface wave is given by

$$t_{i,j}^{(-,\text{app})}\left(t^{(\text{lps})}\right) = t_{i,j}^{(-)} + \delta t_i^{(\text{ins})}\left(t^{(\text{lps})}\right) - \delta t_j^{(\text{ins})}\left(t^{(\text{lps})}\right). \tag{4}$$

Here, $t_{i,j}^{(+)}$ and $t_{i,j}^{(-)}$ are the true arrival times of the direct surface waves, i.e., the direct surface waves in $G(\mathbf{x}_j, \mathbf{x}_i, t)$ and $G(\mathbf{x}_j, \mathbf{x}_i, -t)$, respectively. Consequently, by definition, $t_{i,j}^{(+)} = -t_{i,j}^{(-)}$. It is useful to note that a temporal change in the medium (e.g. Lindner et al., 2018) does not affect the equality between $t_{i,j}^{(+)}$ and $-t_{i,j}^{(-)}$, as it merely modifies the Green's function.

Summing the left-hand and right-hand sides of equations (3) and (4), and subsequently substituting the linear parametrization defined in Equation (2), we find

$$\begin{aligned}\left(t_{i,j}^{(+,\text{app})} + t_{i,j}^{(-,\text{app})}\right)&\left(t^{(\text{lps})}\right) \\ &= 2\delta t_i^{(\text{ins})}\left(t^{(\text{lps})}\right) - 2\delta t_j^{(\text{ins})}\left(t^{(\text{lps})}\right) \\ &= 2a_i t^{(\text{lps})} + 2b_i - 2a_j t^{(\text{lps})} - 2b_j.\end{aligned} \tag{5}$$

The variables here are shown schematically in Figure 1. In the ideal case that (i) the station couple is illuminated uniformly from all angles, (ii) spurious energy has effectively been stacked out in the time-averaging process, and (iii) the recordings are not subject to clock errors and/or drift, the right-hand side of Equation (5) evaluates to zero. If this is the case, then $t_{i,j}^{(+,app)} = -t_{i,j}^{(-,app)} = -t_{i,j}^{(-)} = t_{i,j}^{(+)}$. If, however, the measured $t_{i,j}^{(+,\text{app})}$ and $t_{i,j}^{(-,\text{app})}$ are such that the left-hand side of Equation (5) does not coincide with zero (and the aforementioned conditions are fulfilled), this indicates a clock error at either one or both stations. The associated broken time symmetry is illustrated in Figure 1b.

Assuming the number of lapse cross-correlations $N^{(\text{lps})}$ to coincide for all cross-correlation pairs, $t^{(\text{lps})}$ can be discretized as $t_k^{(\text{lps})}$, where $k = 1, 2, \ldots, N^{(\text{lps})}$. In that case, Equation (5) can be written as

$$t_{i,j,k}^{(+,\text{app})} + t_{i,j,k}^{(-,\text{app})} = 2a_i t_k^{(\text{lps})} + 2b_i - 2a_j t_k^{(\text{lps})} - 2b_j. \tag{6}$$

where the indices k in $t_{i,j,k}^{(+,\text{app})}$ indicate that the arrival times of the direct surface waves are associated with lapse time $t_k^{(\text{lps})}$. The procedure involving the determination of the $t_{i,j,k}^{(+,\text{app})} + t_{i,j,k}^{(-,\text{app})}$ is based on the study by Weemstra et al. (2021), and detailed in Section 2.4. The associated practical implementation is explained in Section 3.2. Finally, it is useful to note that we merely assume the number of lapse cross-correlations per station couple to coincide for notational convenience. In practice, both the number of lapse cross-correlations and their timing (i.e., the values of the $t_k^{(\text{lps})}$) may (and will) vary from one station couple to the other.

2.3 Additional arrival time shifts

Differences in amplitude between the causal and acausal arrivals occur if the noise intensity is larger in one stationary-phase direction than in the other (Stehly et al., 2006). Importantly, a non-uniform illumination pattern may also introduce (small) deviations, or time shifts, from the correct arrival time of the causal and acausal surface waves. We denote these additional time shifts by $\delta t_{i,j,k}^{(\text{src})}$ (the superscript 'src' implies that the time shift is associated with the source distribution). This time shift depends on all three indices since the (noise) illumination pattern usually varies as a function of both time (hence the index k) and station couple (hence the indices i and j). The time dependence of this term is due to the fact that the illumination pattern is usually non-stationary (e.g., Yang and Ritzwoller, 2008; Weemstra et al., 2013). The i, j dependence of this term is explained by the fact that the retrieved causal and acausal direct surface wave responses are associated with opposite stationary-phase regions (e.g., Snieder, 2004; Boschi and Weemstra, 2015). Azimuthal variations of the noise intensity in the two directions along the line connecting a station pair i and j, determine the magnitude of this arrival time shift. We therefore distinguish between $\delta t_{i,j,k}^{(+,\text{src})}$ and $\delta t_{i,j,k}^{(-,\text{src})}$, which represent (illumination related) arrival time shifts of the direct waves at positive (causal) and negative (acausal) time lag(s), respectively. In other words, the illumination-induced (additional) arrival time shifts of the causal and acausal direct surface waves can be expected to differ from each other (Weaver et al., 2009; Froment et al., 2010). We parenthetically note that the medium appears to be slower for a positive $\delta t_{i,j,k}^{(+,\text{src})}$, whereas a positive $\delta t_{i,j,k}^{(-,\text{src})}$ makes the medium appear to be faster than the actual medium.

In addition to the illumination-related arrival time shifts, we account for the presence of spurious energy by defining the additional time shifts $\delta t_{i,j,k}^{(+,\text{spur})}$ and $\delta t_{i,j,k}^{(-,\text{spur})}$, which, similar to $\delta t_{i,j,k}^{(+,\text{src})}$ and $\delta t_{i,j,k}^{(-,\text{src})}$, represent shifts in the arrival times of the causal and acausal direct surface waves, respectively (for details we refer to Weemstra et al., 2021). Including these time shifts in our model, Equation (6) reads:

$$\begin{aligned} t_{i,j,k}^{(+,\text{app})} + t_{i,j,k}^{(-,\text{app})} &= 2a_i t_k^{(\text{lps})} + 2b_i - 2a_j t_k^{(\text{lps})} - 2b_j \\ &\quad + \delta t_{i,j,k}^{(+,\text{src})} + \delta t_{i,j,k}^{(-,\text{src})} \\ &\quad + \delta t_{i,j,k}^{(+,\text{spur})} + \delta t_{i,j,k}^{(-,\text{spur})}. \end{aligned} \tag{7}$$

2.4 Determination of $t_{i,j,k}^{(+,\text{app})} + t_{i,j,k}^{(-,\text{app})}$

As explained in Section 2.2, clock errors manifest themselves by breaking the time-symmetry of the lapse cross-correlations. In order to solve for a large number of a_i and b_i (i.e., to determine clock drift for large OBS arrays), time shifts of individual lapse cross-correlations need to be extracted in an automated fashion.

The $t_{i,j,k}^{(+,\text{app})} + t_{i,j,k}^{(-,\text{app})}$ (for all i, j, and k) are the entries of the data vector $\mathbf{t}^{(\text{app})}$. Our procedure starts by computing *a priori* estimates of these $t_{i,j,k}^{(+,\text{app})} + t_{i,j,k}^{(-,\text{app})}$. This estimate is based on the assumption that, for an individual station couple i, j, the drift accumulated over the interval from $t_1^{(\text{lps})}$ to $t_{N^{(\text{lps})}}^{(\text{lps})}$ is the combined result of a_i and a_j (i.e., that it is linear). Based on the presumed stability of both the medium and the noise illumination, the accumulated drift is estimated by cross-correlating the earliest lapse cross-correlation with the latest lapse cross-correlations: ($C_{i,j}(t, t_1^{(\text{lps})})$ is cross-correlated with $C_{i,j}(t, t_{N^{(\text{lps})}}^{(\text{lps})})$). Assuming the drift to be linear and clock errors to coincide with zero at $t_k^{(\text{lps})} = 0$ then results in the sought-after *a priori* estimates of $t_{i,j,k}^{(+,\text{app})} + t_{i,j,k}^{(-,\text{app})}$, which we denote by $t_{i,j,k}^{(\text{a priori})}$. Note that, as such, the *a priori* estimate of $2b_i - 2b_j$ is assumed to be zero (see Equation 6). Clearly, this is a rather strong assumption. If an initial screening reveals that this assumption is not justified, it may be necessary to combine the procedure here with the procedure described in Section 5 of Weemstra et al. (2021). Finally, it is useful to note that instead of station-couple-specific *a priori* estimates, Weemstra et al. (2021) use station-specific *a priori* estimates to obtain $t_{i,j}^{(\text{a priori})}$ (without an index k because the analysis by Weemstra et al. (2021) does not account for clock drift, but merely allows one to determine time-independent clock errors).

The $t_{i,j,k}^{(\text{a priori})}$ are used to fill an initial estimate of the data vector $\mathbf{t}^{(\text{app})}$. By solving the inverse problem (explained below in Section 2.6), we recover *a priori* estimates of the a_i and b_i. As soon as these estimates are obtained, we apply the procedure described in Section 5 of Weemstra et al. (2021). In summary, this involves determining the time windows in which the causal and acausal direct surface waves are expected using (i) a reference surface wave velocity (which can be station-couple specific), (ii) the station-to-station distance, and (iii) *a priori* estimates of a_i and b_j. Knowing the approximate time windows in which the direct causal and acausal surface waves can be expected, the envelopes of the lapse cross-correlations are subsequently computed. The envelopes are used to determine the arrival time of the direct surface wave (either causal or acausal) with the largest amplitude difference between the top and bottom envelope (denoted by t^{est} in Weemstra et al., 2021). Finally, after interpolating the lapse cross-correlation for a time window (with a length of about one period) centered around the *a priori* estimates of $t_{i,j,k}^{(+,\text{app})}$ and $t_{i,j,k}^{(-,\text{app})}$, and cross-correlating the signals in these two-time windows, the desired measurement $t_{i,j,k}^{(+,\text{app})} + t_{i,j,k}^{(-,\text{app})}$ can be obtained. For a detailed description of the entire process, we refer to

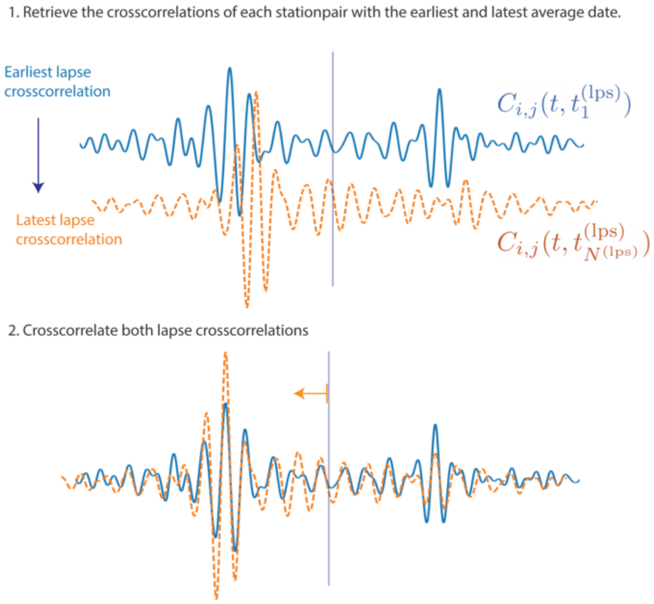


Figure 2 Processing steps for calculating an *a priori* estimate of the combined clock drift of a given station pair.

Weemstra et al. (2021, Section 5).

2.5 Matrix formulation

Assuming we possess synchronous noise recordings by a total of N seismic stations and we compute a total of $N^{(\mathrm{lps})}$ lapse cross-correlations between each station pair, a maximum of $N^{(\mathrm{lps})}$ times $N(N-1)/2$ lapse cross-correlations can be obtained. The set of equations governing the $t_{i,j,k}^{(+,\mathrm{app})} + t_{i,j,k}^{(-,\mathrm{app})}$ can in that case be written as

$$\mathbf{A}\mathbf{t}^{(\mathrm{ins})} + \mathbf{n}^{(\mathrm{src})} + \mathbf{n}^{(\mathrm{spur})} = \mathbf{t}^{(\mathrm{app})}, \tag{8}$$

where the vector $\mathbf{t}^{(\mathrm{ins})}$ contains the sought-for clock drift rates a_i and initial clock errors b_i. This vector has a length of $2N$. The rows of $\mathbf{A}$ relate to different station pairs and lapse times $t_k^{(\mathrm{lps})}$, i.e., they are associated with different $C_{i,j}\left(t, t_k^{(\mathrm{lps})}\right)$. Each column of $\mathbf{A}$ is associated with either an a_i or a b_i. Consequently, $\mathbf{A}$ has dimension $N^{(\mathrm{lps})}(N(N-1)/2) \times 2N$. The length of the vectors $\mathbf{t}^{(\mathrm{app})}$, $\mathbf{n}^{(\mathrm{src})}$, and $\mathbf{n}^{(\mathrm{spur})}$ obviously coincides with the number of rows of $\mathbf{A}$. The vector $\mathbf{t}^{(\mathrm{app})}$ contains the measurements and is often referred to as the 'data vector'. For the sake of clarity, we have detailed these vectors and matrices in Appendix A. Note that throughout this work, both matrices and vectors are indicated in bold; matrices are also capitalized, vectors not.

2.6 Inverting for clock drift

In the model introduced above, we considered the number of lapse cross-correlations $N^{(\mathrm{lps})}$ to coincide for all station pairs. In addition, we assumed these lapse cross-correlations to exist for all possible combinations of stations, i.e., $N(N-1)/2$. In application to field data however, $t^{(+,\mathrm{app})}$ and/or $t^{(-,\mathrm{app})}$ often cannot be determined for all lapse cross-correlations (i.e., all combinations of i, j and k). This implies that the number of rows M of the matrix $\mathbf{A}$ (and hence the number of elements of $\mathbf{t}^{(\mathrm{app})}$, $\mathbf{n}^{(\mathrm{src})}$, and $\mathbf{n}^{(\mathrm{spur})}$) will in practice often be smaller than $N^{(\mathrm{lps})}N(N-1)/2$.

The inability to accurately determine $t^{(+,\mathrm{app})}$ and/or $t^{(-,\mathrm{app})}$ can be due to a number of reasons. First, if two stations are too close to each other with respect to the wavelengths considered, the direct surface-wave response at a positive time will overlap with the direct surface-wave response at a negative time. Second, the absence of sources in one of the two stationary phase directions will prevent the retrieval of the corresponding direct surface-wave response (e.g., Snieder, 2004). Clearly, this also prevents determining the associated arrival time. Third, gaps in the recordings by one or more stations may lead to fewer lapse cross-correlations.

Before we explain the two inversion approaches, we clarify the relation between matrix $\mathbf{A}$ and the ability to obtain a unique (least-squares) estimate of $\mathbf{t}^{(\mathrm{ins})}$. Because, as defined in Appendix A, the rank of $\mathbf{A}$ is two lower than the number of unknowns $2N$ (having a matrix with a rank that is lower than the number of unknowns is often referred to as 'rank deficient'). This indicates that the system of equations is effectively underdetermined. In other words, a unique estimate of $\mathbf{t}^{(\mathrm{ins})}$ does not exist for the system of equations defined in Equation (8). We distinguish between two cases: a land station is included in the network, or no land station is included in the network. In the first case, a unique estimate of $\mathbf{t}^{(\mathrm{ins})}$ exists if a number of conditions are fulfilled. We will detail these in the paragraph below. If the network consists solely of OBSs, a unique estimate of $\mathbf{t}^{(\mathrm{ins})}$ does not exist. We discuss that further below. Finally, an intuitive explanation of the rank deficiency is provided. Consider 10 lapse cross-correlations, each associated with a different $t_k^{(\mathrm{lps})}$ but with the same two OBSs. The matrix $\mathbf{A}$ would be a 10×4 matrix in that case (see Equation 6). Clearly, an infinite number of (least-squares) solutions exist for b_1 and b_2 since adding any (arbitrary) value to both b_1 and b_2 would result in the same left-hand side. In other words, a unique solution for b_1 and b_2 does not exist. The same applies to a_1 and a_2.

In case a station with a UTC-synchronized clock is included in the network (i.e., a land station), the entries of that station can be eliminated from $\mathbf{t}^{(\text{ins})}$ and the associated columns eliminated from $\mathbf{A}$ (see also the discussion in Section 6 and Appendix A in Weemstra et al., 2021). Subsequently, a number of conditions need to be fulfilled for a unique estimate of $\mathbf{t}^{(\text{ins})}$ to exist. First, the system of equations (as defined in Equation 8) needs to contain at least two lines associated with lapse cross-correlations involving that station. These two lapse cross-correlations should be associated with a different lapse time $t_k^{(\text{lps})}$. The land station in the first of the (at least) two lapse cross-correlations may, in fact, be a different land station from the land station associated with the second lapse cross-correlation, as long as the two lapse cross-correlations are associated with different $t_k^{(\text{lps})}$. Second, each of the OBSs needs to be "part of" at least two lapse cross-correlations: there need to be two rows in $\mathbf{A}$ for which the entries associated with that OBS are non-zero. And again, these entries should be associated with different $t_k^{(\text{lps})}$. In case these two conditions are fulfilled, the rank of $\mathbf{A}$ coincides with the number of unknowns ($2N$), and a unique least-squares estimate of the a_i and b_i in $\mathbf{t}^{(\text{ins})}$ exists. Finally, the larger the difference in time between the various lapse cross-correlations of an OBS, the more accurate the estimates of its a_i and b_i.

If the network consists solely of OBSs, a unique estimate of $\mathbf{t}^{(\text{ins})}$ does not exist. In that case, that least-squares estimate of $\mathbf{t}^{(\text{ins})}$ is chosen that has the lowest norm, i.e., that minimizes $||\tilde{\mathbf{t}}^{(\text{ins})}||$, where $\tilde{\mathbf{t}}^{(\text{ins})}$ is any least-squares solution (or *least-squares estimator*) of the underdetermined system of equations. This solution is usually referred to as the *minimum norm* solution. The second condition above, which needed to be fulfilled to obtain a unique estimate of $\mathbf{t}^{(\text{ins})}$, still applies in this case. That is, each of the OBSs still needs to be "part of" at least two lapse cross-correlations. The minimum-norm solution yields an estimator of $\mathbf{t}^{(\text{ins})}$ that allows the OBS recordings to be synchronized with respect to each other, but not with respect to UTC. This is of course, still useful as it would enable tomographic studies using only the OBSs or the localization of seismic events (earthquake hypocenters) below the OBS array.

We consider two estimators of $\mathbf{t}^{(\text{ins})}$. These are the 'ordinary least-squares estimator' $\tilde{\mathbf{t}}^{(\text{ins})}_{(\text{ols})}$, and the 'weighted least-squares estimator' $\tilde{\mathbf{t}}^{(\text{ins})}_{(\text{wls})}$. We refer to Weemstra et al. (2021) for a detailed description (and derivation) of these estimators and will only provide a brief explanation of these two estimators here. The ordinary least-squares estimator minimizes the misfit function $||\mathbf{t}^{(\text{app})} - \mathbf{A}\mathbf{t}^{(\text{ins})}||$ and hence does not account for (potential) variations in the $\delta t_{i,j,k}^{(\text{src})}$ and/or $\delta t_{i,j,k}^{(\text{spur})}$ for different i, j, k. The weighted least-squares estimator, instead, exploits the inverse proportionality of the illumination-related arrival time shifts (i.e., the inverse proportionality of $\delta t_{i,j,k}^{(\text{src})}$) to the true station-to-station travel time $t_{i,j}$ (as derived by Weaver et al., 2009). But since this travel time is usually not known, it uses the station-to-station distances $|\mathbf{x}_j - \mathbf{x}_i|$ as a proxy for the $t_{i,j}$. Measurements (i.e., individual $t_{i,j,k}^{(+,\text{app})} + t_{i,j,k}^{(-,\text{app})}$) associated with lapse cross-correlations between stations (i and j) that are further apart are hence assigned larger weights in the inversion.

3 Implementation & application to data

In this section, we describe the workflow that allows the estimators of $\mathbf{t}^{(\text{ins})}$ to be computed (Section 3.2). Although predominantly methodological aspects of the workflow are discussed (results are presented in Section 4), some examples with field data are shown. We therefore start by introducing the IMAGE's seismic network and its lapse cross-correlations (Section 3.1). After describing the workflow, we dedicate one subsection to our package `OCloC` (Section 3.3). We finish this section with a description of a bootstrapping procedure that allows the stability of the recovered clock drift values to be assessed (Section 3.4).

3.1 The IMAGE data set

For heuristic purposes, the explanation of some processing steps of our workflow includes these steps' application to a set of lapse cross-correlations. These lapse cross-correlations are retrieved from recordings of ambient seismic noise acquired on and around the Reykjanes peninsula, SW Iceland (Jousset et al., 2020a). It concerns lapse cross-correlations between a subset of the stations considered by Weemstra et al. (2021). Specifically, about one year of noise recorded by 30 land stations and 17 OBSs is used (see Figure 3 for the station locations). The OBSs in this experiment are equipped with Seascan clocks (SEASCAN microcomputer compensated crystal oscillators), which are temperature-compensated.

The lapse cross-correlations are computed by averaging individual station-to-station cross-correlations over a 100-day period. These individual cross-correlations are computed per hour with a 50% overlap. We refer to Weemstra et al. (2021) for a detailed description of the computation of the hourly cross-correlations. Averaging individual (hourly) cross-correlations is performed in a two-step process. First, daily cross-correlations are computed based on a maximum of 47 hourly cross-correlations ($24 \times 2 - 1$). Subsequently, these daily cross-correlations are averaged. Importantly, gaps in the recordings by one or both stations are accounted for in the sense that the timing of a lapse cross-correlation, i.e., its $t_k^{(\text{lps})}$, is defined as the average time of the individual cross-correlations. Gaps in the data can cause the average time of the correlations ($t_k^{(\text{lps})}$) to deviate from the center of the 100-day period. Note that the $t_k^{(\text{lps})}$ are allowed to differ between different station couples, as they are explicitly included in $\mathbf{A}$ (see also

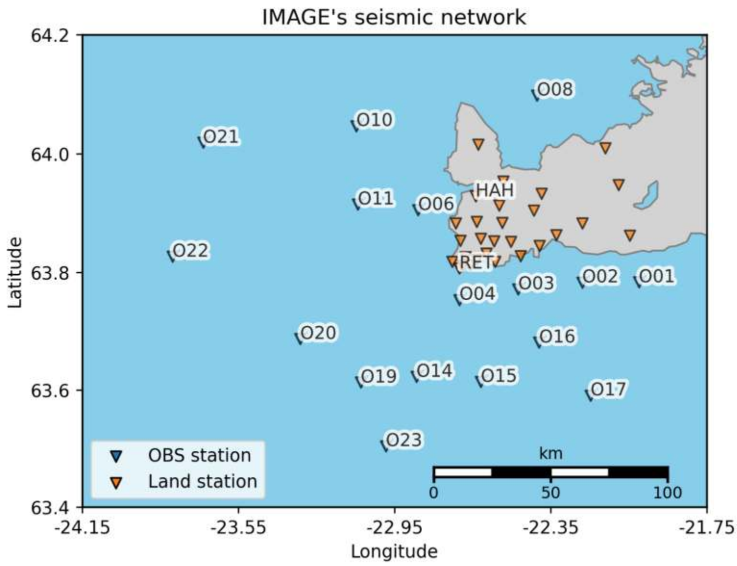


Figure 3 On-and off-shore stations of IMAGE's seismic network, SW Iceland, whose lapse cross-correlations were used in this study. Note that the numbering of the OBSs runs up to 23, whereas only 17 OBSs are included in our set of lapse cross-correlations (some stations did not sample the ambient seismic field sufficiently long and were hence excluded from our analysis; see also Figure S1 in Weemstra et al. (2021)). Only the land stations 'HAH' and 'RET', which are analyzed in Sections 3.2.5 and 4, are labeled due to space constraints.

Appendix A). In case the number of individual cross-correlations contributing to a lapse cross-correlation does not exceed 75% of the maximum number of individual cross-correlations (which is 100×47), that lapse cross-correlation is discarded. An overview of the data availability is given in Figure S1 of Weemstra et al. (2021).

3.2 Workflow

To determine and correct clock drift using lapse cross-correlations of ambient seismic noise, we adopt the processing sequence in Figure 4. It is this workflow that is implemented in `OCloC`. The workflow comprises five steps. We now dedicate one subsection to explain and discuss each of these steps.

3.2.1 Initial screening

In Figure 5a, all stations and ray paths associated with the available lapse cross-correlations are shown. To get a first impression of whether or not the OBS recordings are subject to clock drift, one can plot the different lapse cross-correlations in a single plot (i.e., time-averaged cross-correlations associated with different $t_k^{(\mathrm{lps})}$). In Figure 5b, we depict lapse cross-correlations between stations 020 and HAH (land station) for 5 different lapse times. Potential clock drift of an OBS manifests itself as a shift in time of the lapse cross-correlations: for this specific station couple, the lapse cross-correlations associated with larger $t_k^{(\mathrm{lps})}$ are shifted to a later time.

Prior to the determination of clock drift, it is important to choose an adequate bandpass filter. For the IMAGE data, the surface waves in the retrieved interferometric responses have the highest signal-to-noise ratios (SNRs) between 0.1 and 0.4 Hz. In general, however, the pass band depends on parameters such as the nominal station-to-station distance, the amplitude of the noise sources, the illumination pattern, and the geographical location of the OBS array (e.g., Yang and Ritzwoller, 2008). Note that, due to surface-wave dispersion, lower frequency bands usually result in smaller separations in time of the causal and acausal surface wave peaks.

Processing step	Methods
1. Initial screening	Bandpass filtering
2. Selecting eligible cross-correlations	Selection of SNR and distance thresholds
3. Determination of the $t_{i,j,k}^{(+,\mathrm{app})} + t_{i,j,k}^{(-,\mathrm{app})}$	Filtering out stations without enough station-connections / periods
4. Solving the inverse problem	Computation of the desired least-squares estimator
5. An iterative approach	Improving the estimation of $t_{i,j,k}^{(+,\mathrm{app})} + t_{i,j,k}^{(-,\mathrm{app})}$ by performing iterative inversions
6. Qualitative uncertainty analysis	Bootstrap re-sampling

Figure 4 Workflow for the determination of OBS clock drift using lapse cross-correlations of ambient seismic noise between a large number of OBSs (computed from large-N ocean-bottom seismometer deployments).

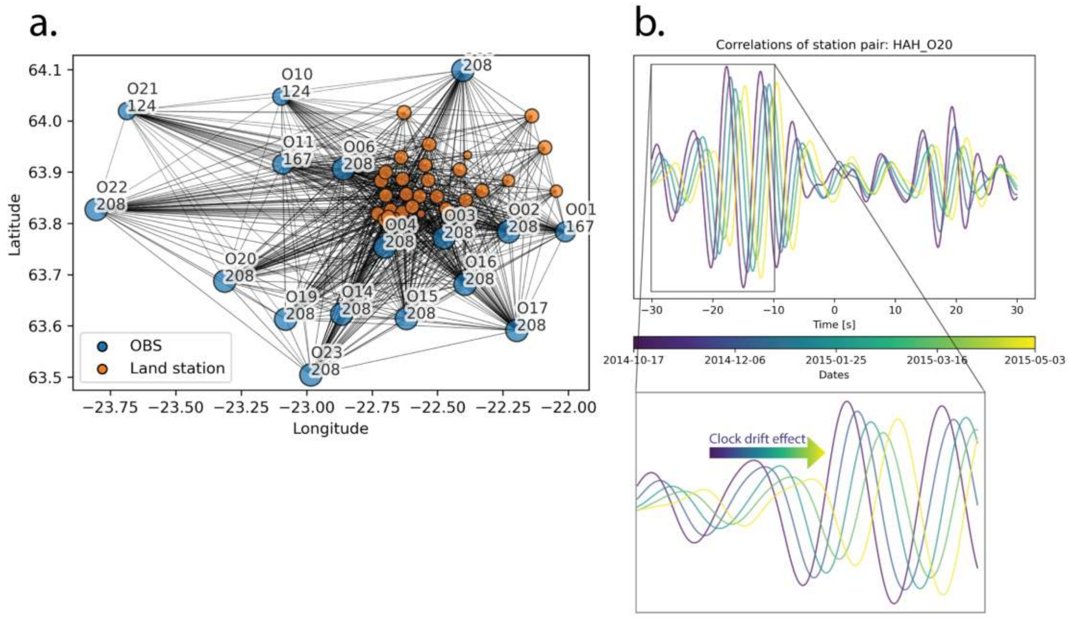


Figure 5 a. All seismic stations and ray paths; blue and orange circles correspond to OBSs and land seismometers, respectively. Only the station names of the OBSs are indicated. Below each station name, the number of available lapse cross-correlations involving that specific station is depicted. b. All lapse cross-correlations for a given station pair. The colors indicate the average timing ($t_k^{(\mathrm{lps})}$) of the lapse cross-correlation.

Importantly, the choice of frequency band also strongly affects the capability to determine the $t_{i,j,k}^{(+,\mathrm{app})} + t_{i,j,k}^{(-,\mathrm{app})}$ of individual lapse cross-correlations.

3.2.2 Selecting eligible lapse cross-correlations

There are two parameters that determine a lapse cross-correlation's eligibility to be included in the clock error estimation process: the SNR threshold and the station-to-station distance threshold. Together, these parameters determine which lapse cross-correlations are included in the inversion and which are not (i.e., whether their $t_{i,j,k}^{(+,\mathrm{app})} + t_{i,j,k}^{(-,\mathrm{app})}$ will be determined and added to data vector $\mathbf{t}^{(\mathrm{app})}$ or not).

In general, the quality of the measurements (i.e., the $t_{i,j,k}^{(+,\mathrm{app})} + t_{i,j,k}^{(-,\mathrm{app})}$) strongly depends on the signal-to-noise ratio (SNR). If the SNR is too low, the algorithm experiences difficulties determining the arrival times of the interferometric responses. Low SNRs are mainly due to low-intensity illumination from (one of) the stationary-phase regions (Snieder, 2004; Weaver et al., 2009). Consequently, the measurements may be inaccurate, or even subject to cycle skipping (Weemstra et al., 2021). Obviously, inaccurate entries in the data vector $\mathbf{t}^{(\mathrm{app})}$ (i.e., inaccurate $t_{i,j,k}^{(+,\mathrm{app})} + t_{i,j,k}^{(-,\mathrm{app})}$) adversely affect the inversion results. A clear example is shown in Figure 6a, where the causal peaks of the lapse cross-correlations between stations O08 and O21 have low SNRs. In this case, the determination of the arrival time of the causal peak is not straightforward and hence may result in inaccurate $t_{i,j,k}^{(+,\mathrm{app})} + t_{i,j,k}^{(-,\mathrm{app})}$. Both SNRs, of the causal and acausal interferometric direct surface waves, need to exceed the SNR threshold for the lapse cross-correlations to be included in the inversion. For details regarding the computation of the SNR, we refer to Weemstra et al. (2021).

The second important parameter when it comes to the accuracy of the $t_{i,j,k}^{(+,\mathrm{app})} + t_{i,j,k}^{(-,\mathrm{app})}$ is the station-to-station distance. If two stations are too close to each other, the direct surface-wave response at a positive time (i.e., the causal arrival) will overlap with the direct surface-wave response at a negative time (i.e., the acausal arrival). Consequently, our algorithm will simply not be able to correctly determine the $t_{i,j,k}^{(+,\mathrm{app})} + t_{i,j,k}^{(-,\mathrm{app})}$ for those station couples. To prevent the inclusion of such measurements in the system of equations, the user must set a station-to-station distance threshold. This threshold is expressed in terms of wavelengths since the ability to distinguish the causal from the acausal arrival does not merely depend on the surface wave travel time, but on the ratio between the travel time and the (dominant) period of the interferometric surface waves. This threshold needs to be set at the start of the workflow (for further details regarding the station-to-station distance threshold we refer to Weemstra et al., 2021). A lapse cross-correlation's station-to-station distance needs to exceed the distance threshold for that lapse cross-correlation to be included in the inversion (i.e., for the $t_{i,j,k}^{(+,\mathrm{app})} + t_{i,j,k}^{(-,\mathrm{app})}$ to be determined and added to data vector $\mathbf{t}^{(\mathrm{app})}$).

Using the IMAGE lapse cross-correlations, we investigate how different thresholds affect the number of eligible lapse cross-correlations. If the thresholds are set too high, there will not be sufficient lapse cross-correlations to (accurately) determine the clock drift of all OBSs (i.e., the vector $\mathbf{t}^{(\mathrm{app})}$ will be relatively short). Conversely, if the thresholds are too low, we add too many inaccurate data points to the data vector, in turn leading to less accurate a_i and b_i (and hence less accu-

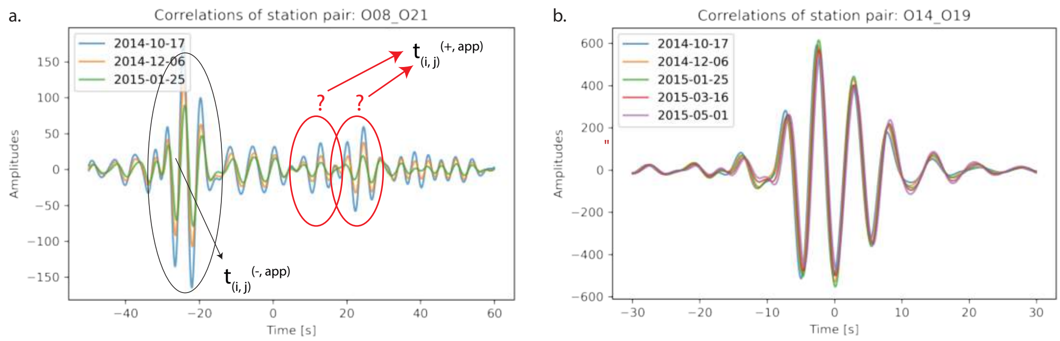


Figure 6 a. Lapse cross-correlations between OBSs O08 and O21: the signal-to-noise ratio of the causal wave is rather low, complicating the determination of $t^{(+,\mathrm{app})}_{i,j,k} + t^{(-,\mathrm{app})}_{i,j,k}$. b. Lapse cross-correlations between stations O14 and O19: The station-to-station distance of these stations is so small (10.6 km) that the causal and acausal surface waves overlap (note that for surface waves with a period of 5 seconds that propagate at 3000 m/s, 10.6 km corresponds to only 2/3 of a wavelength).

rate clock drift estimates). Figure 7 depicts the number of eligible station pairs exceeding a specific combination of thresholds. Obviously, lower thresholds result in a higher number of eligible lapse cross-correlations. Although a higher number of lapse cross-correlations implies a larger number of measurements, it has been shown that station-to-station distance thresholds in the range of 2 to 4 wavelengths and SNR thresholds of about 15 yield the most accurate clock errors (Weemstra et al., 2021). The latter values, however, are based on synthetic data. Here, we therefore choose a slightly more conservative SNR threshold of 30, while setting the station-to-station distance threshold to 2.5. The lapse cross-correlations fulfilling these criteria (i.e., exceeding these thresholds) are added to $\mathbf{t}^{(\mathrm{app})}$ and hence enter the inversion.

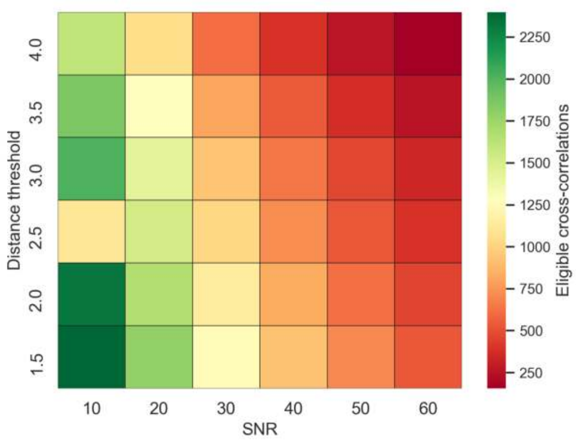


Figure 7 Number of eligible lapse cross-correlations for different station-to-station distances and SNR thresholds.

3.2.3 Determination of the $t^{(+,\mathrm{app})}_{i,j,k} + t^{(-,\mathrm{app})}_{i,j,k}$ for all selected combinations i, j, k

Although the calculation of $t^{(+,\mathrm{app})}_{i,j,k} + t^{(-,\mathrm{app})}_{i,j,k}$ is explained in Section 2.4, a few "practicalities" require attention. First, the algorithm computes the aforementioned *a priori* clock drift estimate only for lapse cross-correlations that exceed the SNR and station-to-station distance thresholds. This may result in some stations having few unique "connections" with other stations. It may be better to, for each station, set both a minimum number of unique connections and a minimum number of total lapse cross-correlations. The lapse cross-correlations, associated with a station that does not exceed these thresholds, will be eliminated from the system of equations (i.e., the data vector $\mathbf{t}^{(\mathrm{app})}$ will be shortened, and the number of rows and columns of the matrix $\mathbf{A}$ decreases).

Second, to recover a unique estimate of a station's clock drift (i.e., of the a_i), that station needs to be associated with lapse cross-correlations at various lapse times $t^{(\mathrm{lps})}_k$ (recall the discussion in Section 2.6). In other words, $t^{(+,app)}_{i,j,k} + t^{(-,app)}_{i,j,k}$ needs to have been determined for various k for that station. By defining (i) a minimum number of correlation periods, (ii) the number of different lapse times required, and (iii) the minimum separation in days between an OBS' lapse cross-correlations, a unique solution can be guaranteed (i.e., provided lapse cross-correlations involving a land station are present, the system of equations will then not be rank deficient). These parameters can be set in `OCloC`.

Finally, a notorious problem in the inversion is what has been referred to as "cycle skipping" by Weemstra et al. (2021). That is, a measurement deviates from the true $t^{(+,\mathrm{app})}_{i,j,k} + t^{(-,\mathrm{app})}_{i,j,k}$ by approximately one period. Needless to say, the inclusion of these measurements in the inversion leads to incorrect a_i and b_i. In Appendix B, we describe a procedure allowing one to detect such outliers and discard them.

3.2.4 Solving the inverse problem

As mentioned in Section 2.6, two inversion strategies can be adopted (both implemented in `OCloC`): the ordinary least squares estimator $\tilde{\mathbf{t}}^{(\mathrm{ins})}_{(\mathrm{ols})}$ and the weighted least-squares estimator $\tilde{\mathbf{t}}^{(\mathrm{ins})}_{(\mathrm{wls})}$ can be computed. The ordinary least-squares estimator can be used if the noise sources uniformly illuminate the stations from all directions. In that case, the vector $\mathbf{n}^{(\mathrm{src})}$ in Equation (8) coincides with $\mathbf{0}$ and the only source of noise is $\mathbf{n}^{(\mathrm{spur})}$. Assuming the entries of the latter vector to have coinciding variance, the ordinary least-squares estimator $\tilde{\mathbf{t}}^{(\mathrm{ins})}_{(\mathrm{ols})}$

will give the most accurate estimate of $\mathbf{t}^{(\mathrm{ins})}$ (in a least-squares sense).

In case the surface wave illumination is not uniform (as is in practice often the case; Yang and Ritzwoller, 2008; Stehly et al., 2006), $\mathbf{n}^{(\mathrm{src})}$ does not coincide with zero, and it is more appropriate to compute the weighted least-squares estimator $\tilde{\mathbf{t}}^{(\mathrm{ins})}_{(\mathrm{wls})}$, where the station-to-station distances $|\mathbf{x}_j - \mathbf{x}_i|$ act as weights (see Section 2.6, and, for further details, Weemstra et al., 2021). In Section 5.1, we demonstrate the superiority of the weighted least-squares estimator, which was previously shown using synthetic noise cross-correlations. Finally, in the absence of lapse cross-correlations with recordings by a land station, the minimum-norm solution is computed. In this case, the recovered b_i differs from the true (unknown) b_i by a common time shift.

3.2.5 An iterative approach

Upon solving the inverse problem using the *a priori* estimates $t^{(\mathrm{a\ priori})}_{i,j,k}$, we obtain an initial estimate of the a_i and b_i values of each station. The latter can subsequently be used to improve the estimation of $t^{(+,\mathrm{app})}_{i,j,k} + t^{(-,\mathrm{app})}_{i,j,k}$ as they can be used to predict the arrival time of the interferometric surface wave responses (see also Weemstra et al., 2021). It is therefore recommended to perform several inversions, each iteration using the previously obtained a_i and b_i to guide the estimation of the $t^{(+,\mathrm{app})}_{i,j,k} + t^{(-,\mathrm{app})}_{i,j,k}$ resulting in an updated data vector $\mathbf{t}^{(\mathrm{app})}$, until the recovered a_i and b_i do not change anymore. By simply plotting the evolution of the recovered a_i and b_i, it is possible to determine when this is the case.

By plotting all lapse cross-correlations associated with a single station couple in one frame, and doing this separately for the corrected (using the obtained a_i and b_i) and uncorrected set of lapse cross-correlations, a (qualitative) impression of the result is obtained. If the lapse cross-correlations associated with different lapse times align, then the clock drift is successfully removed. An example of a successful clock drift removal is shown in Figure 8. It is clear that the lapse cross-correlations suffered from clock drift of OBS O20 (Figure 8a). Once the clock drift is removed, the lapse cross-correlations associated with different lapse times nicely align, as shown in Figure 8b.

3.3 `OCloC`

The methodology presented in this paper has been implemented in `OCloC`. In particular, `OCloC` allows the workflow detailed in the previous subsection to be executed. `OCloC` is an open-source Python package that has been tested for the operating systems Linux and macOS. We chose Python as `OCloC`'s main programming language for its open-source, versatile, and cross-platform compatible nature, which is widely used in the Earth sciences (e.g., Werthmüller et al., 2021; Rücker et al., 2017). In the case of `OCloC`, the portability of Python enabled us to outsource specific computational aspects to a pre-compiled Fortran module.

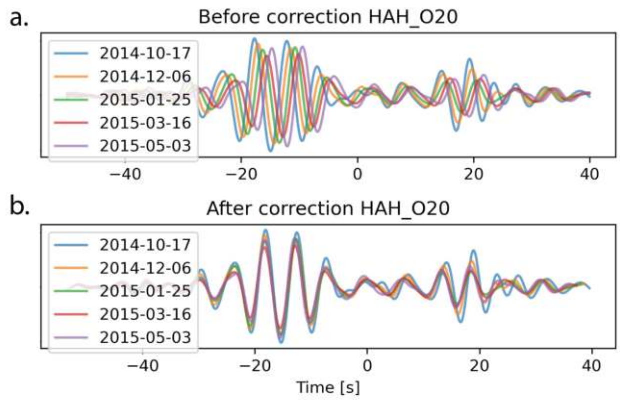


Figure 8 Lapse cross-correlations between land station HAH and OBS O20 for different lapse times. a. Cross-correlations before applying time corrections b. Lapse cross-correlations after correcting the clock drift of O20 using the a_i and b_i recovered by means of the iterative weighted least-squares inversion.

Through the application of seismic interferometry, the proposed correction of clock errors is contingent on the availability of synchronous noise recordings. The computation of the lapse cross-correlations, however, is deliberately left out of `OCloC`. The reason is that it will be nearly impossible to account for the plethora of different (pre-)processing approaches (Seats et al., 2012; Groos et al., 2012; Weemstra et al., 2014; Fichtner, 2014). This implies that users of the package have complete freedom regarding pre-processing (e.g., one-bit normalization, spectral whitening, etc.) and potential filter settings while computing the lapse cross-correlations, and that they are expected to do this themselves prior to the application of `OCloC`. The lapse cross-correlations can subsequently be imported as `OCloC` objects.

`OCloC`'s functionality includes loading lapse cross-correlation files, storing and accessing station metadata, and solving the linear systems of equations in Section 2.5 in a (weighted) least-squares sense. It also has some other supporting functions. To keep the use of `OCloC` simple, a hierarchical object-oriented design has been adopted. This kind of architecture breaks down the whole process of determining and correcting clock errors into solvable chunks while letting the user know when an error occurred and how to prevent it.

The main object types of `OCloC` are: `ClockDrift`, `ProcessingParameters`, `Correlation`, and `Station`. Figure 9 depicts, schematically, the algorithm's object hierarchy. These objects need some clarification:

1. `ClockDrift`: The outermost layer of the hierarchical structure. The user deals with this object for the main processing steps described in Section 3.2. This object stores the different `Station` and `Correlation` objects in the form of lists. This object also stores the system of equations, described in Section 2, in the form of a `Pandas` dataframe (Wes McKinney, 2010a). The different methods of `ClockDrift` provide access to correlation files, station metadata, plotting functions, and different processing tools required for the algorithm's usage.

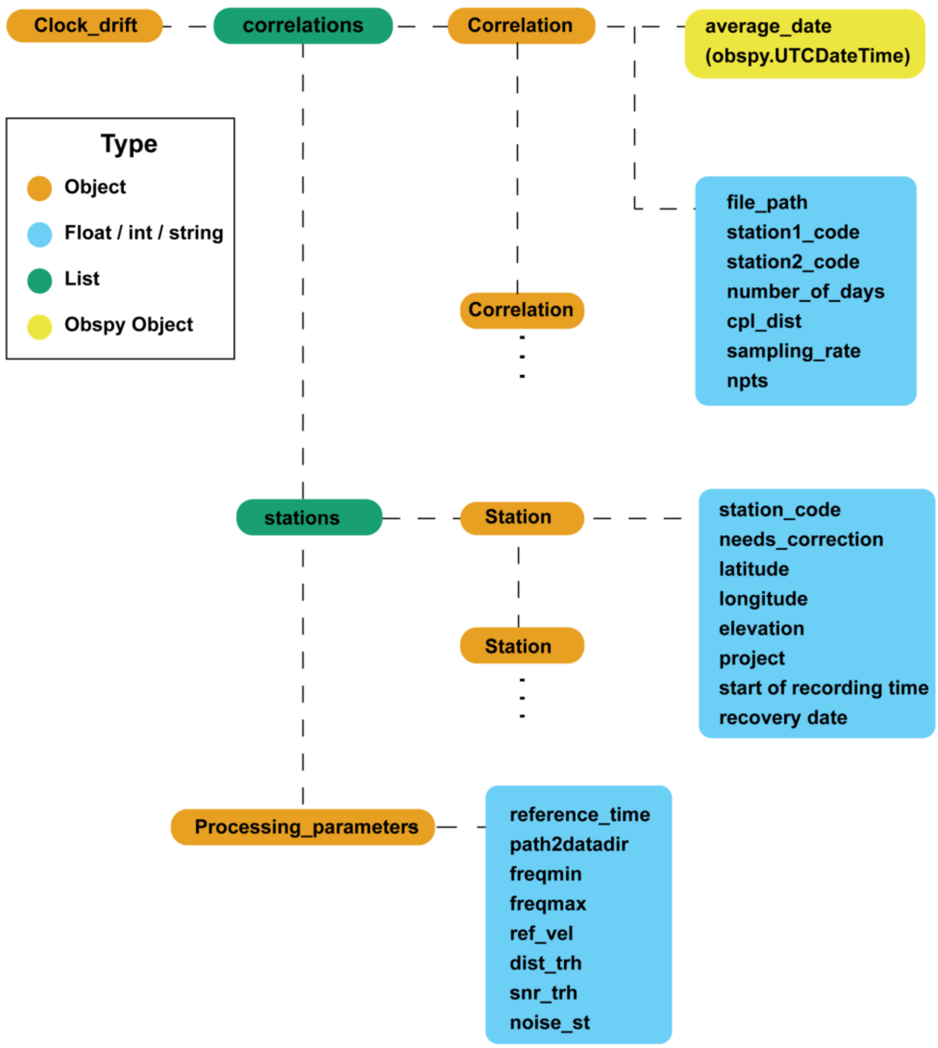


Figure 9 Object hierarchy of `OCloC`.

2. `ProcessingParameters`: The recovery of clock errors depends on the adequate selection of some pre-processing steps. `ProcessingParameters` object stores the value of these parameters. These parameters are the band-pass filter's corner frequencies, the SNR threshold, and the station-to-station distance threshold. These parameters are detailed in Section 3.2.

3. `Correlation`: Stores the metadata of each cross-correlation file such as the station names, the lapse time $t_k^{(lps)}$, and station-to-station distance, among others. Additionally, this object has functions to compute $t_{i,j,k}^{(+,\text{app})} + t_{i,j,k}^{(-,\text{app})}$, together with the signal-to-noise ratios of the causal and acausal surface wave arrivals.

4. `Station`: Contains metadata such as location, code, and timestamp when the station started recording. Moreover, after solving the linear system of equations, the recovered clock errors, i.e., the a_i and b_i, can be retrieved through these objects.

In addition to the core module, `OCloC` incorporates third-party dependencies that yield advanced functionality, namely, the `Numpy` programming library (Harris et al., 2020), several signal processing functions from `Obspy` (Krischer et al., 2015), and the data visualization tools of `Pandas` (Wes McKinney, 2010b) and `Matplotlib` (Hunter, 2007). For specific details regarding the package installation and usage, please refer to the online documentation available at https://ocloc.readthedocs.io.

3.4 Bootstrap re-sampling

To verify the robustness of the obtained results, we repeat the inversion several times using different sets of measurements $t_{i,j,k}^{(+,\text{app})} + t_{i,j,k}^{(-,\text{app})}$. By repeating the inversion multiple times, mean values and confidence intervals of the sought-after parameters are obtained. One way to artificially generate different sets of measurements is using bootstrap re-sampling (Efron, 1982). Bootstrapping is a statistical method that falls under the broader class of re-sampling methods. It allows one to estimate statistical properties of interest such as sample averages and variances (Schnaidt and Heinson, 2015). Effectively, it gives an indication of which results are likely and which are less likely without computing new lapse cross-correlations. Here, we seek to obtain an estimate of the variance of the estimated a_i and b_i. Instead of using all the measured data points, we sample with replacement (Efron, 1982, 1992). In practice, we generate a large number of data vectors $\mathbf{t}^{(\text{app})}$ (usually referred to as 'realizations'), each with the same length as the original data vector, but with values that are drawn from the original data vector, allowing duplicates. Specifically, we perform the following steps:

I. An initial estimate of clock drift is obtained following steps one to five of Figure 4. A SNR threshold of 30 and a station-to-station distance threshold of 2.5 wavelengths are applied. It is necessary to check that the recovered a_i and b_i values are no longer changing after several iterations. This results in the data vector $\mathbf{t}^{(\text{app})}$ that serves as the input of our bootstrapping procedure.

II. Allow sampling with replacement by randomly selecting measurements of $t_{i,j}^{(+,\text{app})} + t_{i,j}^{(-,\text{app})}$ (bootstrapped samples).

III. Once having re-sampled the measurements, perform the inversion and store the recovered a_i^* and b_i^* values of each station.

IV. Repeat steps II and III one thousand times. By doing so, we store 1000 possible realizations of the recovered a_i^* and b_i^* values.

V. Based on all the a_i^* and b_i^* realizations, compute a statistical measure, such as 95% confidence intervals, for each of the stations.

To identify stations with relatively uncertain a and b values (either due to a limited number of data points, or due to a lot of noise on the lapse cross-correlations associated with that specific station), we estimate the standard deviation and 95% confidence intervals (CI) from the 1000 realizations. The CI represents the range in which 95% of the a_i^* and b_i^* values lie.

The bootstrap approach allows one to identify OBSs with narrow or large confidence intervals. Narrower confidence intervals suggest the recovered a and b are well-determined, whereas stations with larger confidence intervals point to larger uncertainties in the recovered clock errors. In the absence of noise, i.e., $\mathbf{n}^{(\text{src})}$ and $\mathbf{n}^{(\text{spur})}$ both coinciding with $\mathbf{0}$, all a_i^* and b_i^* of a given station should coincide and hence the 95% confidence interval would be zero.

4 Results

4.1 Clock drift rates

We computed the weighted least-squares estimator of $\mathbf{t}^{(\text{ins})}$ for the OBSs in the IMAGE's network. Our findings indicate that all OBS stations experienced clock drift. Compared to the other OBSs, the clock drift of OBS O20 was particularly large. Table 1 summarizes the estimated clock drift rates (i.e., the a_i) and incurred clock errors at the time of deployment. The latter may deviate slightly from the b_i because the b_i represents the clock errors on August 21, 2014 ($t^{(\text{lps})} = 0$), whereas most stations were not deployed exactly on that date. In addition, we list the measured skews in the last column. To compare these skew values, we provide in the fifth column the clock error at the time of recovery computed using the estimated a_i and b_i. Note that most OBS recordings end prior to that date due to full disks. We also obtained a drift estimate for OBS O21, which had no skew value documented due to a dead battery at the time of recovery. The incurred initial clock errors at the time of deployment ranged from a minimum of -0.404 s to a maximum of 0.037 s.

The bootstrap re-sampling introduced in Section 3.4 allows us to estimate the variance of the recovered a_i and b_i. By generating 1000 different data vectors (realizations) and subsequently performing a separate inversion for each of the generated data vectors, 1000 weighted least-squares estimators of $\mathbf{t}^{(\text{ins})}$) are obtained. The standard deviation of the recovered a_i^* from the a_i (recovered using the original $\mathbf{t}^{(\text{app})}$) is listed in column 3 of Table 1. In Figure 10, we visualize the recovered a_i and b_i, including the bootstrap-derived uncertainties.

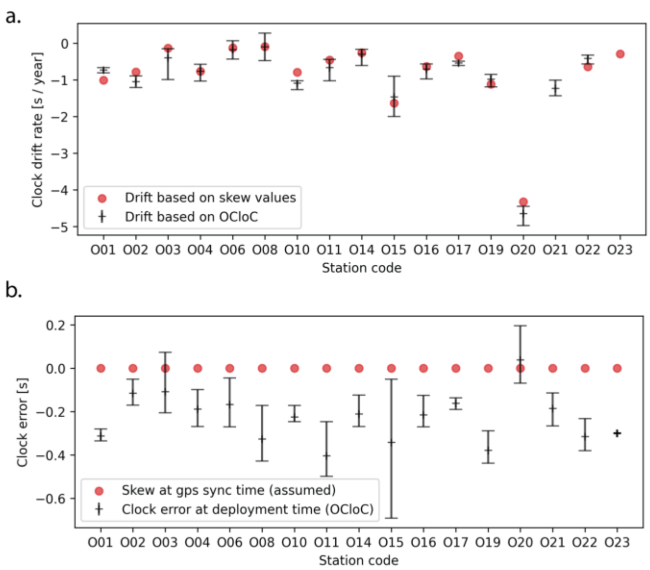


Figure 10 Comparison between clock drifts obtained in this study, and the measured skews. a. Comparison of clock drift rates estimated based on the skew values and the total recording time (red dots) and the a_i obtained from our weighted least-squares inversion (black crosses). The error bars correspond to the 95% confidence intervals resulting from the bootstrap re-sampling. Note that no skew value was reported for OBS O21 as this instrument's battery died before recovery. b. Comparison of the initial clock error at the OBS' deployment time. In both a and b, OBS O23 has no error bars as this OBS was associated with too few data points in $\mathbf{t}^{(\text{app})}$ to be successfully included in the bootstrapping procedure.

For all station pairs, we plotted the waveform (mis)alignment of the different lapse cross-correlations to verify the effective removal of the clock errors. Figure 11 shows the time-lapse cross-correlations between OBS O01 and land station RET (which is devoid of clock errors) in three states: a) uncorrected, b) corrected

Table 1 Estimated clock drift rates (a_i) of the OBSs (2nd column), and their corresponding standard deviation (3rd column). The clock drift rates in column four are based on the measured skew values, assuming a linear drift and no clock errors at $t^{(\text{lps})} = 0$. The estimated clock errors at deployment and recovery time in the fifth and sixth columns, respectively, are computed by substituting the estimated a_i and b_i in Equation (2) with $t^{(\text{lps})}$ set to each OBS' day of deployment and recovery. OBS O21 had no skew value reported, as the battery died before recovery. Station O23 has no standard deviation because it was associated with too few data points in $\mathbf{t}^{(\text{app})}$ to be successfully included in the bootstrapping procedure. This is probably due to the relatively low SNRs of the lapse cross-correlations involving this station.

Station name	Clock drift rate based on OCloC [s/year]	σ [s/year]	Clock drift rate based on skew values [s/year]	Clock error at deployment time (OCloC) [s]	Clock error at recovery time [s]	Measured skew [s]
O01	-0.734739	0.042760	-1.011740	-0.312892	-1.055899	-1.023125
O02	-1.055136	0.097011	-0.782087	-0.115908	-1.182942	-0.790906
O03	-0.401807	0.251438	-0.136396	-0.109015	-0.515270	-0.137906
O04	-0.770560	0.153252	-0.765888	-0.189565	-0.968726	-0.774437
O06	-0.172589	0.152249	-0.126476	-0.167177	-0.343849	-0.129468
O08	-0.104288	0.245437	-0.096861	-0.326683	-0.433947	-0.099625
O10	-1.095582	0.073045	-0.789557	-0.225825	-1.354066	-0.813093
O11	-0.667440	0.179920	-0.457559	-0.404513	-1.089599	-0.469656
O14	-0.304885	0.147440	-0.326255	-0.211671	-0.462262	-0.268156
O15	-1.465134	0.340357	-1.633493	-0.342048	-1.839114	-1.669093
O16	-0.712585	0.131879	-0.635126	-0.216183	-0.944088	-0.648781
O17	-0.547051	0.034074	-0.350884	-0.161655	-0.720530	-0.358468
O19	-0.985476	0.115891	-1.119642	-0.378849	-1.388873	-1.147531
O20	-4.652023	0.176057	-4.324439	0.038416	-4.744142	-4.445781
O21	-1.234367	0.141809	N/A	-0.185787	-1.456213	N/A
O22	-0.415065	0.077167	-0.643822	-0.315925	-0.743071	-0.662562
O23	-0.312865	N/A	-0.289709	-0.300249	-0.620852	-0.296875

using skew-derived drift rates, and c) corrected using OCloC's weighted least-squares estimates of the drift rates and the initial clock errors. Before correction, the later lapse cross-correlations shift monotonically to an earlier time (Figure 11a). The skew-derived corrections shift the lapse cross-correlations to later times. In this case, however, the skew-derived drift rate appears to "overcorrect" the lapse cross-correlations: later lapse cross-correlations now shift monotonically to a later time (Figure 11b). Finally, shifting the lapse cross-correlations using the weighted least-squares inversion for the a_i and b_i results in lapse cross-correlations that properly align (Figure 11c).

In Figure 11, the skew-derived drift and the drift recovered using the weighted least-squares inversion are compared for a single station couple only. In Figure 12, a more systematic and quantitative comparison of the linear drift based on our code ('OCloC-drift') and the skew values ('skew-drift') is presented for three OBSs. The drifts of all the other OBSs are shown in Appendix C. Figure 12 also shows the time offsets between the lapse cross-correlations and a reference lapse cross-correlation (RCF). We only use the cross-correlations between the OBSs and land stations. For each station pair, the highest signal-to-noise ratio cross-correlation is selected as the RCF. The time offsets correspond to the time shift that maximizes the Pearson correlation coefficient between the RCF and each cross-correlation. The skew-derived drift in the top figures assumes that there is no initial clock error at the onset of deployment (i.e., $b = 0$), and all subsequent time offsets are linearly interpolated based on this assumption. For the bottom figures, the time offset at deployment time corresponds to

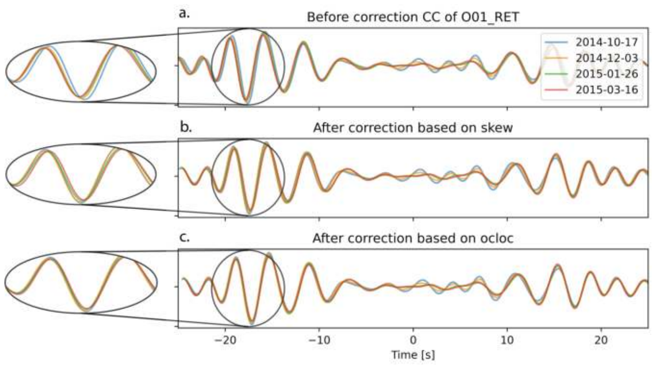


Figure 11 Lapse cross-correlations between OBS O01 and land-station RET in the frequency range of 0.2 to 0.4 Hz. Colors indicate the average time of all time windows that contribute to the lapse cross-correlation, and this color scheme is consistent across all three sub-figures (legend provided in the upper right corner of figure a). a. Original lapse cross-correlations prior to any corrections. b. Lapse cross-correlations after clock drift correction using skew values. c. Lapse cross-correlations after clock drift correction using the (OCloC-derived) drift rates (a_i) and initial clock errors (the b_i) estimated in this study.

the OBS' b value (or initial clock error) estimated from the weighted least-squares inversion (again, all subsequent offset times are interpolated accordingly).

The OCloC-drift corrections of OBS01 and OBS02 (Figures 12a and 12b) seem to align better with the time offset between the cross-correlations and the RCF. This is not the case for OBS10 (Figure 12c), where the skew-based clock drift aligns better with the time offsets. Station O10, however, is also one of the stations with the shortest recording time, which results in fewer time-lapses. This highlights one limitation of our approach: the need for longer monitoring time to include more lapse cross-correlations, because that provides tighter constraints for the inversion process. In Section 5.3, we further discuss the implications of the OBSs not monitoring for a full year.

4.2 Comparing inversion strategies

The fact that a non-uniform illumination pattern can break the time symmetry of the retrieved surface-wave responses is detrimental to the method presented in this study. Weemstra et al. (2021) showed that applying a weighted least-squares inversion based on station-to-station distances decreases the adverse effects of a non-uniform surface wave illumination. Using synthetic recordings of ambient seismic noise, these authors demonstrated the advantage of the weighted least-squares estimate over the ordinary least-squares estimate. To evaluate the accuracy of the weighted least-squares estimator $\tilde{\mathbf{t}}^{(\text{ins})}_{(\text{wls})}$ in the presence of a non-uniform surface wave illumination, we compare it to the ordinary least-squares estimator $\tilde{\mathbf{t}}^{(\text{ins})}_{(\text{ols})}$. To do so, we used the bootstrap re-sampling approach introduced in Section 3.4. Figure 13 shows the histogram and cumulative distribution of 1000 bootstrap realizations of the a^* values of both inverse strategies. We used the same starting parameters and data vectors in both cases. The weighted inversions are shown in red, whereas the ordinary least-squares inversions are shown in blue. Figure 13a shows the distribution of the bootstrap realizations for all stations centered around 0 (mean values have been subtracted for each OBS individually). Figure 13b shows the cumulative distribution of the bootstrap realizations, with the 5th and 95th quantiles marked as vertical lines. The weighted least-squares distribution has narrower confidence intervals than the ordinary least-squares distribution.

For data vectors associated with large station-to-station distance thresholds, we do not find significant differences between both inversion strategies. This is expected because the threshold removes measurements associated with station couples that are closer to each other, which hence removes those lapse cross-correlations that are susceptible to larger illumination-related noise errors. For data vectors resulting from decreasing station-to-station distance thresholds, however, we find that the weighted inversion results in narrower bootstrap confidence intervals.

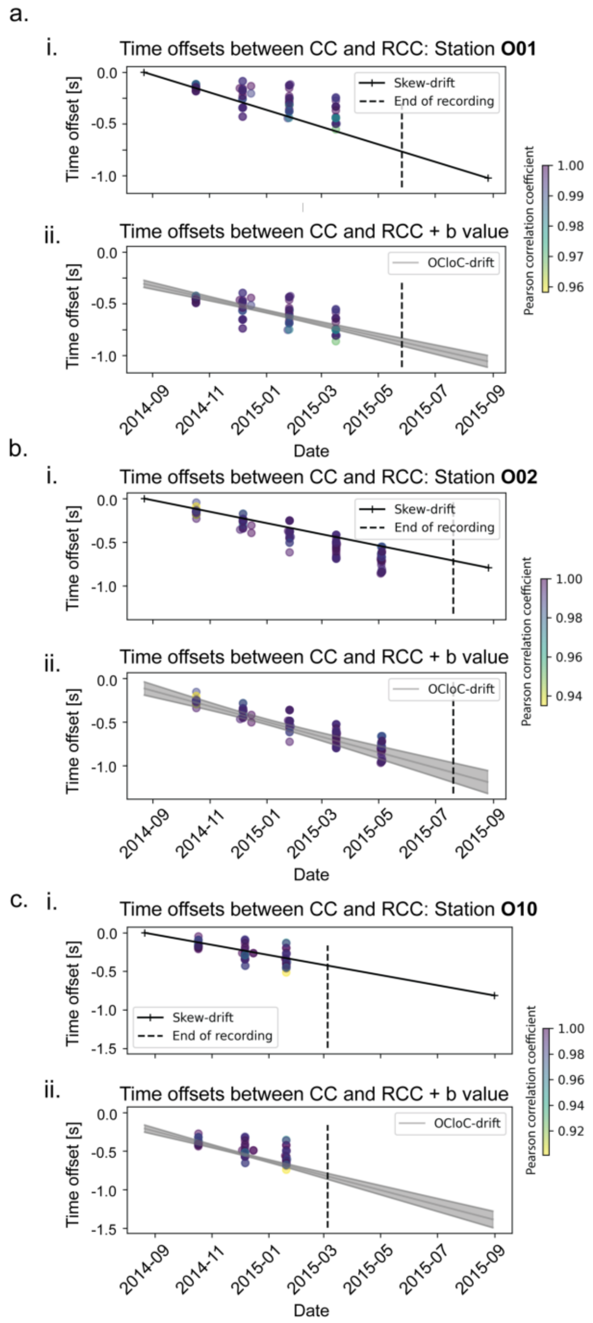


Figure 12 Comparison between (i) the skew-derived linear clock drift and (ii) the linear clock drift recovered using the weighted least-squares inversion for three selected OBSs. **Top**: Time offsets between cross-correlations and a reference cross-correlation (RCF) assuming no initial clock error at the onset of deployment. **Bottom**: Time offsets considering the initial clock error (b value) at deployment time. The drift based on our code (weighted least-squares inversion) and the confidence intervals are dubbed 'OCloC-drift', while the drift based on the skew values is termed 'skew-drift'. The highest signal-to-noise ratio cross-correlation for each station pair is chosen as the RCF. The depicted time offsets result from maximizing the Pearson correlation coefficient between the RCF and the other lapse cross-correlations, plus a correction based on the b value to the skew correction.

5 Discussion

5.1 The effect of the surface wave (noise) illumination pattern

A limitation of the presented method is the fact that a non-uniform illumination pattern can lead to deviations of the retrieved surface-wave responses from the true surface-wave responses (e.g., Tsai, 2009; Weaver et al., 2009). As such, timing errors due to a non-uniform illumination pattern (captured in $\mathbf{n}^{(\text{src})}$) lead

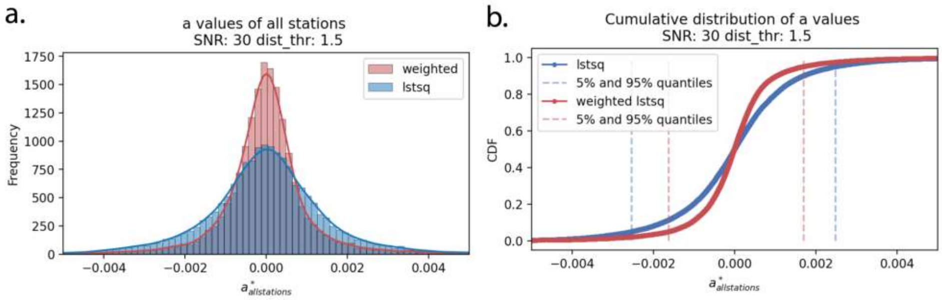


Figure 13 Bootstrap analysis of a^* values for all stations showing the 1000 realizations stacked for all stations. The mean values have been removed. The results of the weighted least-squares inversion are shown in red, and the results of the ordinary least-squares inversion are depicted in blue. a. Frequency histogram of the recovered values for all stations, with the probability density function overlaid. b. Cumulative distribution of the bootstrap realizations, with the 5th and 95th quantiles marked.

to deviations of the recovered drift from the true clock drift. Consequently, bootstrap confidence intervals can be expected to be larger for more pronounced non-uniform illumination patterns. The bootstrapping results presented in Section 4.2 show that the distance-based weighted least-squares inversion result in both a lower spread of the distribution of the a^* values (Figure 13a) and a narrower range between the 5th and 95th quantiles (Figure 13b).

The bootstrapping results presented in Section 4.2 confirm the earlier, synthetic-data-based findings by Weemstra et al. (2021). Compared to the ordinary least-squares inversion, the weighted least-squares inversion decreases the adverse effects of a non-uniform illumination pattern. Note that the reasoning above can also be turned around: the fact that the weighted least-squares inversion results in more accurate clock drift estimates strongly suggests a (time-varying) non-uniform surface wave illumination. Given the available literature (Stehly et al., 2006; Mulargia, 2012; Weemstra et al., 2013) in general, and the large differences in SNRs between (some of) the causal and acausal direct surface waves in particular, this can hardly be surprising.

5.2 Validation using only land stations

We run a separate test using only the lapse cross-correlations between the land stations. Lapse cross-correlations involving OBSs are discarded. Apart from two stations, we pretend these land stations to be suffering from clock errors and hence neither eliminate the columns associated with any of them from $\mathbf{A}$ (in reality, those stations' a and b coincide with zero of course) nor any of its two entries from $\mathbf{t}^{(\text{ins})}$. We subsequently compute the weighted least-squares estimator of $\mathbf{t}^{(\text{ins})}$. The inversion yields drift rates (i.e., a_i) of maximum 0.1 s/year, which demonstrates that (i) noise on the data (i.e., non-zero $\mathbf{n}^{(\text{src})}$ and $\mathbf{n}^{(\text{spur})}$) prevents the recovery of drift rates of 0 s/day and (ii) that drift rates lower than 10^{-4} s/day cannot be recovered unambiguously (for our specific station configuration, noise illumination, and frequency band). The maximum b_i that is recovered has a value of 0.12 s, but this is an outlier in the sense that for most of the stations, the estimated initial clock error at $t^{(\text{lps})} = 0$ does not exceed 0.05 s. Although the recovered a_i and b_i do not coincide with zero, we know that, in practice, the land stations do not suffer from clock drift and/or initial clock errors. Effectively, this experiment tells us that our approach allows us to successfully recover a seismic station's clock drift with an uncertainty of approximately 0.1 s/year.

5.3 On the validity of the assumption of linear clock drift and an initial clock error

While introducing our model (Section 2.2), we assumed the clock drift rates to be constant. Specifically, we formulated a time-dependent clock error $\delta t_i^{(\text{ins})}(t^{(\text{lps})})$ which drifts at a constant rate a_i, while allowing for a possible clock error b_i at $t^{(\text{lps})} = 0$. The latter is introduced to allow for an initial clock error at deployment time. This could, for example, be invoked by the temperature shock while the OBS is sunk (Zhang et al., 2023). We discuss in this section (i) the differences between the skew-derived drift rates and the recovered drift rates (i.e., the a_i), (ii) the fact that the b_i are non-zero, and (iii) the relation between these two observations.

First, we would like to emphasize that the differences between the skew-derived drift rates and the parameters recovered using the weighted least-squares inversion (i.e., the a_i and b_i) yield clock errors at the time of recovery that differ at maximum 0.62 s (compare the last two columns of Table 1); most of them much less. This suggests that the skew values are rather representative of the clock drift of the OBSs at the time of recovery (see also Figure 10). Furthermore, Figure 12 and Appendix C preclude a decision as to which drift rate is“better” based on the time offsets of the consecutive lapse cross-correlations.

Second, Figure 12 and Appendix C also do not allow us to draw firm conclusions regarding the estimated initial clock errors at the onset of deployment (i.e., the b_i). Upon comparison with the time offsets between the individual lapse cross-correlations and the RCF, however, the clock errors estimated using our weighted least-squares inversion seem to be slightly more accurate than the skew-derived clock errors for most OBSs. This would confirm the existence of a (non-zero) initial clock error. For OBS O01, for example, the waveform alignment presented in Figure 11 strongly supports the estimated initial clock error (b) of -0.31 s. In particular because the estimated clock error at the time of recovery almost coincides with the skew for this station (see Table 1). By not taking into account the initial clock error for this OBS, the skew-based corrections effectively overcompensate the observed clock drift. This is evident from the shift of later lapse cross-correlations to positive times in Figure 11b. In contrast, `OCloC`-based corrections do not yield any (visible) residual drift (Figure 11c). This implies that an initial clock error at the time of deployment (i.e., a non-zero b) is indeed needed to explain the observed clock errors. Note that OBS O01 is used as an example because this station has one of the largest initial clock errors at deployment time, whereas its estimated clock error at recovery time almost coincides with the measured skew.

The fact that the `OCloC`-corrected lapse cross-correlations align better than the skew-corrected lapse cross-correlations can hardly be surprising. The drift rate estimates and the initial clock errors at the time of deployment are based on these very lapse cross-correlations. Therefore, it can be misleading to conclude from this observation that our approach yields more accurate drift rates than the skew-derived drift rates. This is because although the weighted least-squares inversion mitigates the effect of arrival time shifts resulting from a non-uniform surface wave illumination, it will not undo it entirely. Illumination-related arrival time shifts (i.e., non-zero $\delta t_{i,j,k}^{(+,\mathrm{src})}$) may still have some effect. However, given the fact that (i) we averaged hourly cross-correlations over a period of 100 days, (ii) an SNR-threshold of 30 was imposed, and (iii) a station-to-station separation threshold of 2.5 wavelengths needed to be exceeded, we do not expect that these illumination-related arrival time shifts to be the cause of b_i as high as 0.3 or 0.4 s. The experiment discussed in Section 5.2 supports this claim.

Considering the above, we identify two possible explanations for the fact that the initial clock errors at the time of deployment are found to be non-zero and have values as high as (minus) 0.4 s. One explanation is that they result from the temperature shock during the OBS' descent to the ocean floor (see e.g., Zhang et al., 2023). In other words, they are real. This would not be surprising considering the experimental results by Gardner and Collins (2012), who find that the drift rates of the SEASCAN clocks may change significantly in the weeks after a temperature shock (in practice: after deployment). A second possible explanation for their deviation from zero stems from the fact that the OBSs experience seasonal temperature variations during their deployment at (relatively) shallow depths. The study by Jochumsen et al. (2016), for example, reports on seasonal seawater-temperature variations on the order of five degrees centigrade at those depths. This is consistent with the temperature variations within the data logger, which reveal annual temperature variations of about four degrees centigrade (these temperature sensors have a resolution of one degree only). In general, an OBS' drift rate is temperature dependent (Shariat-Panahi et al., 2009). However, we do not expect the drift rate of the SEASCAN clocks to suffer from such temperature variations: the SEASCAN clocks are temperature compensated (Gardner and Collins, 2012). Nonetheless, if such a seasonally varying drift would exist, it may be more appropriate to have our drift model (Equation 2) include a sinusoid with a period of one year. This may be the topic of future work.

Of all recovered drift rates, the drift rate by OBS O20 stands out (see Figure 10). This may well be explained by the fact that, compared to the other OBSs, the logger and hence SEASCAN crystal oscillator of OBS O20 was newer. It was only two years old at the time of deployment, whereas the loggers (and hence clocks) of the other OBSs were approximately 8 years old (Alfred Wegener Institute, personal communication, 2023). This matters because of a natural process in the crystal oscillator, which is referred to as aging. Aging implies that the drift rate of an oscillator slowly changes with time. Essentially, it is the time derivative of the drift rate (Gardner and Collins, 2012). The aging of the crystal is a very important factor when it comes to the drift rate of the SEASCAN clocks, with younger crystals usually aging faster. And even though aging can be mitigated by regular recalibration of the SEASCAN clocks, it could well have been the cause of the larger drift rate of the SEASCAN clock of OBS O20.

5.4 Performance in the absence of land stations?

OBS arrays in remote oceanic regions will not have the benefit of land stations in their near vicinity. In that case, lapse cross-correlations between the OBSs and a station with a correct (UTC) reference time do not contribute to the data vector $\mathbf{t}^{(\mathrm{app})}$. The system of equations will, in that case, be underdetermined (the rank of $\mathbf{A}$ being lower than the number of unknowns) and that weighted least-squares estimator $\tilde{\mathbf{t}}_{(\mathrm{wls})}^{(\mathrm{ins})}$ is chosen that has the lowest norm (see Section 2.6). The minimum-norm solution yields an estimator of $\mathbf{t}^{(\mathrm{ins})}$ that allows the OBS recordings to be synchronized with respect to each other, but not with respect to UTC. In other words, the recovered b_i differs from the true (unknown) b_i by a common time shift, but the drift rates (i.e., the a_i) can still be recovered (with some uncertainty, of course). This is still useful as it would enable tomographic studies using only the OBSs, or the localization of seismic events (earthquake hypocenters) below the OBS array.

The accuracy of the recovered clock drift parameters depends linearly on the wave frequency. That is, lapse cross-correlations at higher frequencies will hence result in more accurate estimates of clock drift than lapse cross-correlations at lower frequencies, provided the illumination pattern and the SNRs at both frequencies coincide. In practice, the latter is often not the case: lapse cross-correlations at lower frequencies usually benefit from more uniform noise illumination patterns (e.g., Yang and Ritzwoller, 2008). It may therefore be beneficial to include measurements associated with different frequency bands in $\mathbf{t}^{(\mathrm{app})}$. It is beyond the scope of this work to investigate this here.

5.5 Which projects can benefit from OCloC?

There are several methods that can be used for correcting OBS clock errors. The fastest to implement is simply using the recovered skew values and assuming a linear drift rate. Here, however, we show that this method may not be reliable. Moreover, it may not be possible because the battery has died before recovery. Other methods require correcting each OBS one by one by simply evaluating cross-correlations of ambient seismic noise in a non-automated manner. This requires a level of inspection that is not attractive (time-wise) for large-N OBS arrays. OCloC is suitable for such type of deployment as it automatically and simultaneously computes clock drift rates of all OBSs. Other cases where a GPS clock is lost, particularly with only on-land-station deployments, can significantly benefit from OCloC.

Projects that do not benefit from our approach are those with a limited deployment time. The reason is that OCloC requires the retrieval of interferometric surface wave responses at positive and negative times. In addition, lapse cross-correlations need to be computed at different lapse times $t_k^{(\mathrm{lps})}$ (at least two). To retrieve both responses, noise cross-correlations need to be averaged over a sufficiently long time. Here "sufficiently long" is location, processing, and frequency dependent (e.g., Yang and Ritzwoller, 2008; Seats et al., 2012; Snieder, 2004). In our case (Reykjanes peninsula, spectral whiting prior to cross-correlation, and 0.2–0.4 Hz frequency band), individual noise cross-correlations were averaged over 100 days to obtain surface waves with sufficiently high SNRs at both positive (causal peak) and negative (acausal peak) time.

For projects that might not be suitable for OCloC, alternative solutions exist, such as the methodologies proposed by Sens-Schönfelder (2008); Hable et al. (2018); Loviknes et al. (2020); Jousset et al. (2013); Gouédard et al. (2014), among others.

6 Conclusions

We introduced a new method to recover, simultaneously, clock drift rates of large numbers of ocean bottom seismometers. Our approach relies predominantly on the time-symmetry of the retrieved interferometric surface wave responses, but also includes the perceived temporal stability of the lapse cross-correlations in the workflow (see Appendix B). Contrary to existing approaches, our method also (i.e., in addition to the drift rate) allows one to recover an initial clock error at the time of deployment. Two situations can be distinguished: OBS deployments including land stations and OBS deployments without stations devoid of a clock error. Drift rates will successfully be recovered in both situations. The absolute time, however, will be meaningless in case no land station (or another station devoid of clock errors) is "connected" to OBS deployment by means of a number of lapse cross-correlations. Results can be analyzed using a qualitative uncertainty analysis via bootstrap re-sampling. Finally, the presented methodology is implemented in OCloC, an accessible Python package with an object-oriented design.

We test OCloC using the seismic noise data acquired during IMAGE's seismic campaign in and around the Reykjanes Peninsula (Iceland). We find that all OBSs in the network suffered from clock drift. In particular, we find that the skew did not allow accurate recovery of the OBSs' drift rates. Using our approach, it was possible to detect the OBSs' initial clock error at the time of deployment. Finally, we showed that a weighted least-squares inversion, where receiver pairs are weighted by station-to-station distances, significantly reduces errors caused by deviations of the noise illumination pattern from uniform.

7 Acknowledgments

This work received funding from the Competitive Research Grant ZAFRAN: OSR-CRG2019-4076 from KAUST granted to Sigurjón Jónsson. The data used in this project was provided by the IMAGE project (Integrated Methods for Advanced Geothermal Exploration) and can be retrieved from http://www.image-fp7.fr/Pages/default.aspx. Instruments for the IMAGE project were provided by the GIPP (Geophysical Instrument Pool Potsdam) and the DEPAS (German Instrument Pool for Amphibian Seismology). In particular, we would like to thank Mechita Schmidt-Aursch from the Alfred Wegener Institute for her comprehensive answers to our many questions regarding the mechanism dictating the DEPAS OBS' clock drift rate. DN has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 956965. We thank Guus Hoogewerf for his additions to the code and the analyses performed in the scope of his M.Sc. thesis. Similarly, we thank Janneke de Laat for previously contributing to the Fortran routines used in this study. Finally, we would like to express our sincere gratitude to the reviewers of this manuscript. Their questions and comments have certainly improved our study.

Code availability

Name of the package: OCloC (OBS Clock Correction)
Contact: d.f.naranjohernandez@tudelft.nl
Program language: Python
The documentation can be found at: https://ocloc.readthedocs.io
The source codes are available for download at: https://github.com/davidn182/ocloc

Competing interests

The authors have no competing interests.

References

Aki, K. Space and time spectra of stationary stochastic waves, with special reference to microtremors. *Bulletin of the Earthquake Research Institute*, 35:415–456, 1957. doi: 10.15083/0000033938.

Blanck, H., Jousset, P., Hersir, G. P., Ágústsson, K., and Ólafur G. Flóvenz. Analysis of 2014–2015 on- and off-shore passive seismic data on the Reykjanes Peninsula, SW Iceland. *Journal of Volcanology and Geothermal Research*, 391:106548, 2020. doi: 10.1016/j.jvolgeores.2019.02.001.

Boschi, L. and Weemstra, C. Stationary-phase integrals in the cross correlation of ambient noise. *Reviews of Geophysics*, 53 (2):411–451, 2015. doi: 10.1002/2014RG000455.

Claerbout, J. F. Synthesis of a layered medium from its acoustic transmission response. *Geophysics*, 33(2):264–269, 1968. doi: 10.1190/1.1439927.

Dongmo Wamba, M., Montagner, J.-P., and Romanowicz, B. Imaging deep-mantle plumbing beneath La Réunion and Comores hot spots: Vertical plume conduits and horizontal ponding zones. *Science Advances*, 9(4):eade3723, Jan. 2023. doi: 10.1126/sciadv.ade3723.

Draganov, D., Wapenaar, K., Mulder, W., Singer, J., and Verdel, A. Retrieval of reflections from seismic background-noise measurements. *Geophysical Research Letters*, 34(4), 2007. doi: 10.1029/2006GL028735.

Efron, B. *The jackknife, the bootstrap and other resampling plans*. SIAM, 1982. doi: 10.1137/1.9781611970319.

Efron, B. *Bootstrap Methods: Another Look at the Jackknife*. Springer New York, New York, NY, 1992. doi: 10.1007/978-1-4612-4380-9_41.

Fichtner, A. Source and processing effects on noise correlations. *Geophysical Journal International*, 197(3):1527–1531, 2014. doi: 10.1093/gji/ggu093.

Froment, B., Campillo, M., Roux, P., Gouédard, P., Verdel, A., and Weaver, R. Estimation of the effect of nonisotropically distributed energy on the apparent arrival time in correlations. *Geophysics*, 75, 09 2010. doi: 10.1190/1.3483102.

Gardner, A. T. and Collins, J. A. Advancements in high-performance timing for long term underwater experiments: A comparison of chip scale atomic clocks to traditional microprocessor-compensated crystal oscillators. In *2012 Oceans*, pages 1–8, 2012. doi: 10.1109/OCEANS.2012.6404847.

Geissler, W. H., Matias, L., Stich, D., Carrilho, F., Jokat, W., Monna, S., IbenBrahim, A., Mancilla, F., Gutscher, M.-A., Sallarès, V., and Zitellini, N. Focal mechanisms for sub-crustal earthquakes in the Gulf of Cadiz from a dense OBS deployment. *Geophysical Research Letters*, 37(18), 2010. doi: 10.1029/2010GL044289.

Gouédard, P., Seher, T., McGuire, J., Collins, J., and Hilst, R. D. Correction of Ocean-Bottom Seismometer Instrumental Clock Errors Using Ambient Seismic Noise. *Bulletin of the Seismological Society of America*, 104:1276–1288, 2014. doi: 10.1785/0120130157.

Groos, J., Bussat, S., and Ritter, J. Performance of different processing schemes in seismic noise cross-correlations. *Geophysical Journal International*, 188(2):498–512, feb 2012. doi: 10.1111/j.1365-246X.2011.05288.x.

Hable, S., Sigloch, K., Barruol, G., Stähler, S. C., and Hadziioannou, C. Clock errors in land and ocean bottom seismograms: high-accuracy estimates from multiple-component noise cross-correlations. *Geophysical Journal International*, 214 (3):2014–2034, 06 2018. doi: 10.1093/gji/ggy236.

Halliday, D. and Curtis, A. Seismic interferometry, surface waves and source distribution. *Geophysical Journal International*, 175 (3):1067–1087, 12 2008. doi: 10.1111/j.1365-246X.2008.03918.x.

Hannemann, K., Krueger, F., and Dahm, T. Measuring of clock drift rates and static time offsets of ocean bottom stations by means of ambient noise. *Geophysical Journal International*, 196: 1034–1042, 02 2014. doi: 10.1093/gji/ggt434.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., Fernández del Río, J., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E. Array programming with NumPy. *Nature*, 585:357–362, 2020. doi: 10.1038/s41586-020-2649-2.

Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Computing in Science Engineering*, 9(3):90–95, 2007. doi: 10.1109/MCSE.2007.55.

Jochumsen, K., Schnurr, S. M., and Quadfasel, D. Bottom temperature and salinity distribution and its variability around Iceland. *Deep Sea Research Part I: Oceanographic Research Papers*, 111: 79–90, 2016. doi: 10.1016/j.dsr.2016.02.009.

Jousset, P., Budi-Santoso, A., Jolly, A. D., Boichu, M., Surono, Dwiyono, S., Sumarti, S., Hidayati, S., and Thierry, P. Signs of magma ascent in LP and VLP seismic events and link to degassing: An example from the 2010 explosive eruption at Merapi volcano, Indonesia. *Journal of Volcanology and Geothermal Research*, 261: 171–192, 2013. doi: 10.1016/j.jvolgeores.2013.03.014.

Jousset, P., Blanck, H., Franke, S., Metz, M., Águstsson, K., Verdel, A., Ryberg, T., Hersir, G., Weemstra, K., Bruhn, D., and Flovenz, O. "Seismic Tomography in Reykjanes , SW Iceland". In *European geothermal congress 2016*, pages 1–10, 2016. http://europeangeothermalcongress.eu/.

Jousset, P., Hersir, G. P., Blanck, H., Kirk, H., Erbas, K., Hensch, M., Franke, S., Bruhn, D., and Huenges, E. IMAGE (Integrated Methods for Advanced Geothermal Exploration). GFZ Data Services. Other/Seismic Network. 2020a. doi: 10.14470/9Y7569325908.

Jousset, P., Mortensen, A. K., Ómar Fridleifsson, G., Ágústsson, K., and Gudmundsson, M. T. Reykjanes, Iceland: Structure and dynamics of mid-oceanic ridge geo/hydrothermal systems. *Journal of Volcanology and Geothermal Research*, 391:106692, 2020b. doi: 10.1016/j.jvolgeores.2019.106692.

Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C., and Wassermann, J. ObsPy: A bridge for seismology into the scientific Python ecosystem. *Computational Science and Discovery*, 8:014003, 05 2015. doi: 10.1088/1749-4699/8/1/014003.

Lindner, F., Weemstra, C., Walter, F., and Hadziioannou, C. Towards monitoring the englacial fracture state using virtual-reflector seismology. *Geophysical Journal International*, 214(2):825–844, 04 2018. doi: 10.1093/gji/ggy156.

Lobkis, O. and Weaver, R. On the emergence of the Green's function in the correlations of a diffuse field. *Journal of the Acoustical Society of America*, 110(6):3011–3017, 2001. doi: 10.1121/1.1417528.

Loviknes, K., Jeddi, Z., Ottemöller, L., and Barreyre, T. When Clocks Are Not Working: OBS Time Correction. *Seismological Research Letters*, 91(4):2247–2258, 05 2020. doi: 10.1785/0220190342.

Matsumoto, H., Zampolli, M., Haralabus, G., Stanley, J., Mattila, J., and Meral Ozel, N. Interpretation of detections of volcanic activity at Ioto Island obtained from in situ seismometers and remote hydrophones of the International Monitoring System. *Scientific Reports*, 9:19519, 12 2019. doi: 10.1038/s41598-019-55918-w.

Mulargia, F. The seismic noise wavefield is not diffuse. *The Journal of the Acoustical Society of America*, 131(4):2853, 2012. doi: 10.1121/1.3689551.

Rücker, C., Günther, T., and Wagner, F. M. pyGIMLi: An open-source library for modelling and inversion in geophysics. *Computers and Geosciences*, 109:106–123, 2017. doi: 10.1016/j.cageo.2017.07.011.

Schnaidt, S. and Heinson, G. Bootstrap resampling as a tool for uncertainty analysis in 2-D magnetotelluric inversion modelling. *Geophysical Journal International*, 203(1):92–106, 08 2015. doi: 10.1093/gji/ggv264.

Seats, K. J., Lawrence, J. F., and Prieto, G. A. Improved ambient noise correlation functions using Welch's method. *Geophysical Journal International*, 188(2):513–523, 2012. doi: 10.1111/j.1365-246X.2011.05263.x.

Sens-Schönfelder, C. Synchronizing seismic networks with ambient noise. *Geophysical Journal International*, 174(3):966–970, 09 2008. doi: 10.1111/j.1365-246X.2008.03842.x.

Shapiro, N. M. and Campillo, M. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise. *Geophysical Research Letters*, 31(7), 2004. doi: 10.1029/2004GL019491.

Shariat-Panahi, S., Alegria, F. C., Làzaro, A. M., and del Rio, J. Time Drift of Ocean Bottom Seismometers (Obs). In *Proceedings of the 19th IMEKO World Congress on Fundamental and Applied Metrology*, pages 2548–2553, 2009.

Snieder, R. Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase. *Phys. Rev. E*, 69:046610, Apr 2004. doi: 10.1103/PhysRevE.69.046610.

Stehly, L., Campillo, M., and Shapiro, N. M. A study of the seismic noise from its long-range correlation properties. *Journal of Geophysical Research: Solid Earth*, 111(B10), 2006. doi: 10.1029/2005JB004237.

Sánchez-Pastor, P., Obermann, A., Schimmel, M., Weemstra, C., Verdel, A., and Jousset, P. Short- and Long-Term Variations in the Reykjanes Geothermal Reservoir From Seismic Noise Interferometry. *Geophysical Research Letters*, 46(11):5788–5798, 2019. doi: 10.1029/2019GL082352.

Tary, J. B., Géli, L., Henry, P., Natalin, B., Gasperini, L., Çomoğlu, M., Çağatay, N., and Bardainne, T. Sea-Bottom Observations from the Western Escarpment of the Sea of Marmara. *Bulletin of the Seismological Society of America*, 101(2):775–791, 04 2011. doi: 10.1785/0120100014.

Tsai, V. C. On establishing the accuracy of noise tomography travel-time measurements in a realistic medium. *Geophysical Journal International*, 178(3):1555–1564, 2009. doi: 10.1111/j.1365-246X.2009.04239.x.

Wapenaar, K. and Fokkema, J. Green's function representations for seismic interferometry. *GEOPHYSICS*, 71(4):SI33–SI46, 2006. doi: 10.1190/1.2213955.

Warner, M. and Guasch, L. Adaptive Waveform Inversion - FWI Without Cycle Skipping - Theory. In *76th Conference and Exhibition, EAGE, Expanded Abstracts*. European Association of Geoscientists and Engineers, 2014. doi: 10.3997/2214-4609.20141092.

Weaver, R., Froment, B., and Campillo, M. On the correlation of non-isotropically distributed ballistic scalar diffuse waves. *The Journal of the Acoustical Society of America*, 126(4):1817–1826, 2009. doi: 10.1121/1.3203359.

Weemstra, C., Boschi, L., Goertz, A., and Artman, B. Seismic attenuation from recordings of ambient noise. *GEOPHYSICS*, 78(1): Q1–Q14, 2013. doi: 10.1190/geo2012-0132.1.

Weemstra, C., Westra, W., Snieder, R., and Boschi, L. On estimating attenuation from the amplitude of the spectrally whitened ambient seismic field. *Geophysical Journal International*, 197: 1770–1788, 2014. doi: 10.1093/gji/ggu088.

Weemstra, C., de Laat, J. I., Verdel, A., and Smets, P. Systematic recovery of instrumental timing and phase errors using interferometric surface-waves retrieved from large-N seismic arrays. *Geophysical Journal International*, 224(2):1028–1055, 10 2021. doi: 10.1093/gji/ggaa504.

Werthmüller, D., Rochlitz, R., Castillo-Reyes, O., and Heagy, L. Towards an open-source landscape for 3-D CSEM modelling. *Geophysical Journal International*, 227(1):644–659, 06 2021. doi: 10.1093/gji/ggab238.

Wes McKinney. Data Structures for Statistical Computing in Python. In Stéfan van der Walt and Jarrod Millman, editors, *Proceedings of the 9th Python in Science Conference*, pages 56 – 61, 2010a. doi: 10.25080/Majora-92bf1922-00a.

Wes McKinney. Data Structures for Statistical Computing in Python. In Stéfan van der Walt and Jarrod Millman, editors, *Proceedings of the 9th Python in Science Conference*, pages 56 – 61, 2010b. doi: 10.25080/Majora-92bf1922-00a.

Yang, Y. and Ritzwoller, M. H. Characteristics of ambient seismic noise as a source for surface wave tomography. *Geochemistry, Geophysics, Geosystems*, 9(2), 2008. doi: 10.1029/2007GC001814.

Zhang, J., Zhao, M., Yao, Y., and Qiu, X. Analysis of clock drift based on a three-dimensional controlled-source ocean bottom seismometer experiment. *Geophysical Prospecting*, 2023. doi: 10.1111/1365-2478.13384.

Appendices

A Matrix formulation

To clarify the rather mathematical description of the inverse problem, let's consider the following example. If one would compute monthly time-averaged cross-correlations for an OBS deployment of 10 stations that would last one full year, $N^{(\mathrm{lps})}$ would be 12 and N (obviously) 10. This would imply the number of rows of the matrix $\mathbf{A}$ (and the length of the vectors $\mathbf{t}^{(\mathrm{app})}$, $\mathbf{n}^{(\mathrm{src})}$, and $\mathbf{n}^{(\mathrm{spur})}$) would coincide with $12 \times (10 \times 9)/2 = 540$. The length of $\mathbf{t}^{(\mathrm{ins})}$ would coincide with 20 (2×10) and so would the number of columns of $\mathbf{A}$. Expressing then $t_k^{(\mathrm{lps})}$ in terms of days (instead of seconds, which is the customary unit of time) and setting it to zero at the onset of the OBS deployment, this would imply $t_1^{(\mathrm{lps})} \approx 15$, $t_2^{(\mathrm{lps})} \approx 46$, and so on, and so forth.

For N stations, vector $\mathbf{t}^{(\mathrm{ins})}$ can be written as:

$$\mathbf{t}^{(\mathrm{ins})} \equiv \begin{pmatrix} a_1 \\ b_1 \\ a_2 \\ b_2 \\ \vdots \\ a_N \\ b_N \end{pmatrix}, \tag{9}$$

To aid in the interpretation of Equation (8), we depict below (Figure 14) the rows associated with the first lapse cross-correlations (i.e., the lapse cross-correlations associated with $t_1^{(\mathrm{lps})}$) are shown in light blue. In addition, we have depicted in purple (for $t_1^{(\mathrm{lps})}$ only) the elements of the matrix associated with the lapse cross-correlation between stations 1 and 2, and in yellow the elements of the matrix associated with the lapse cross-correlation between stations 2 and 3. Note that, as it stands, the matrix in Figure 14 is rank deficient. This implies that the system of equations is underdetermined, and a unique solution does not exist. If one of the 10 stations is a land station devoid of clock errors, the two columns associated with that station could be eliminated from $\mathbf{A}$ (that OBS' a and b would coincide with zero), and its two entries eliminated from $\mathbf{t}^{(\text{ins})}$. The resulting matrix $\mathbf{A}$ would be full rank, and a unique estimator of $\mathbf{t}^{(\text{ins})}$ would exist.

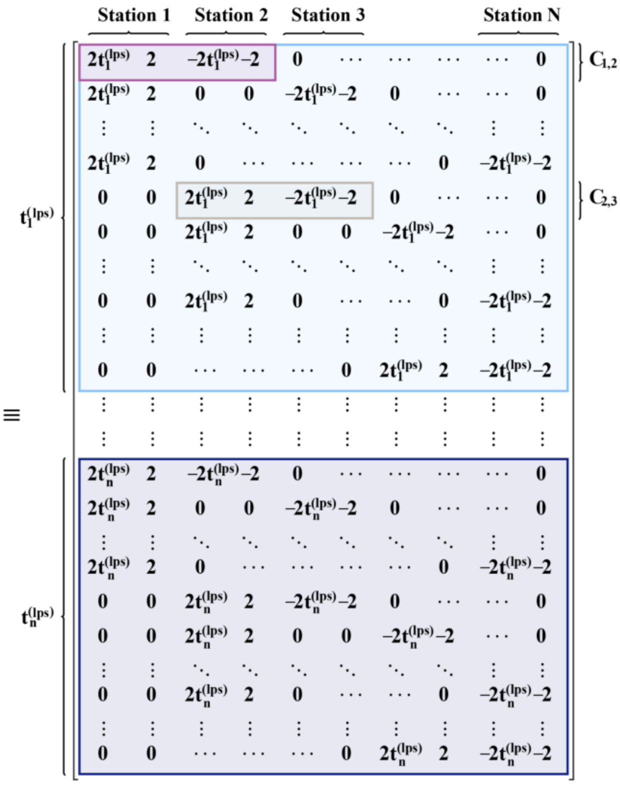


Figure 14 Example of matrix $\mathbf{A}$ when using N stations and n lapse times

B Detection of outliers

When measuring the $t^{(+,\text{app})}_{i,j,k} + t^{(-,\text{app})}_{i,j,k}$ it might be possible to get an erroneous measurement due to a phenomenon that is similar in nature to what is referred to as 'cycle skipping' in full-waveform inversion (e.g., Warner and Guasch, 2014). That is, the measured $t^{(+,\text{app})}_{i,j,k} + t^{(-,\text{app})}_{i,j,k}$ deviates from the "true value" by approximately one period (see also Weemstra et al., 2021). Needless to say, inclusion of these measurements in the inversion leads to incorrect a_i and b_i. To prevent such measurements, we implemented a method that compares the measured $t^{(+,\text{app})}_{i,j,k} + t^{(-,\text{app})}_{i,j,k}$ with the expected $t^{(+,\text{app})}_{i,j,k} + t^{(-,\text{app})}_{i,j,k}$. The latter is computed using the a's and b's recovered during a first inversion. After identifying the outliers, i.e., points that do not follow the overall trend (blue areas in Figure 15), we set a certain threshold for removing or keeping measurements. Repeating this process multiple times allows us to"clean" the data vector from such measurements.

C Clock drifts of each OBS station

In this appendix section, we provide an extended comparison of linear corrections from our code against the

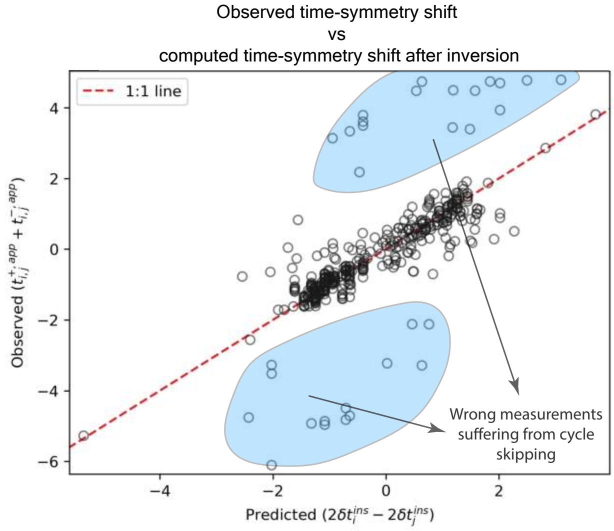


Figure 15 Observed time symmetry shifts plotted against the estimated time symmetry shifts after inversion. The clusters in blue might indicate inaccurate measurements product of cycle skipping.

skew values for each OBS. The figures display the time offsets between the cross-correlations and a chosen reference cross-correlation, offering a detailed view of our approach's alignment with standard skew value corrections across different OBSs. Stations with very high uncertainty (for example O03, O06, and O08) yielded fewer data points ($t^{+,app}_{i,j,k} + t^{-,app}_{i,j,k}$) as their SNR and distance separation did not meet the required thresholds (see Figure 6 for examples of those cross-correlations).

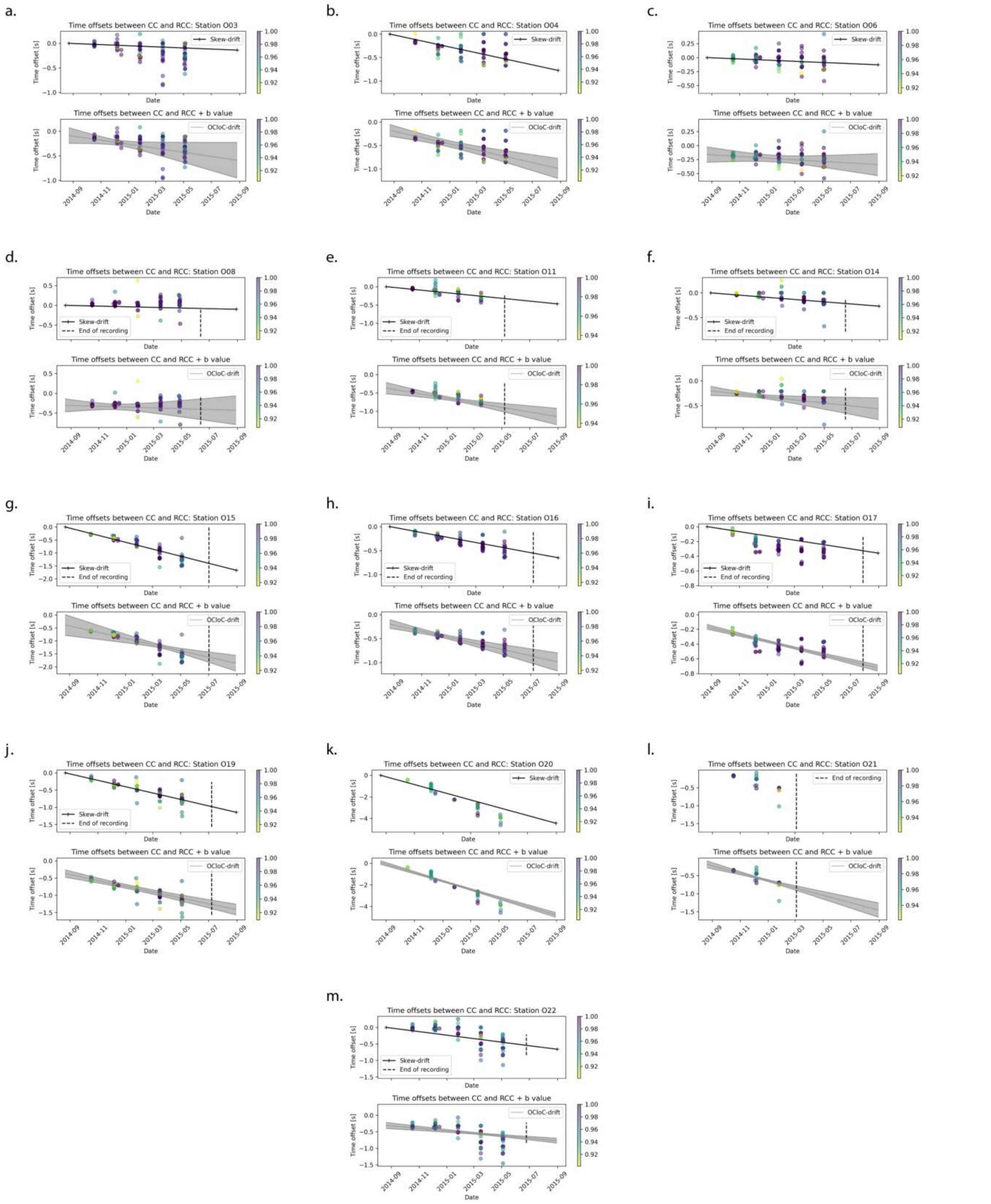


Figure 16 Comparison between the observed clock drift, (i) the skew-derived linear clock drift, and (ii) the linear clock drift recovered using the weighted least-squares inversion of each OBSs (except O01, O02, and O10, which are in Section 4). **Top**: Time offsets between cross-correlations and a reference cross-correlation (RCF) assuming no initial clock error at the onset of deployment. **Bottom**: Time offsets considering the initial clock error (b value) at deployment time. The drift based on our code (weighted least-squares inversion) and the confidence intervals is dubbed '`OCloC`-drift', while the drift based on the skew values is termed 'skew- drift'. The highest signal-to-noise ratio cross-correlation for each station pair is chosen as the RCF. The depicted time offsets result from maximizing the Pearson correlation coefficient between the RCF and the other lapse cross-correlations, plus a correction based on the value of b in the skew correction.

REPORT
doi:10.26443/seismica.v3i1.1154

SEISMICA

Seismoacoustic measurements of the OSIRIS-REx re-entry with an off-grid Raspberry PiShake

Benjamin Fernando * [1], **Constantinos Charalambous** [2], **Christelle Saliby** [3], **Eleanor K. Sansom** [4], **Carene Larmat** [5], **David Buttsworth** [6], **Daniel C. Hicks** [7], **Roy Johnson** [8], **Kevin Lewis** [1], **Meaghan McCleary** [9], **Giuseppe Petricca** [10], **Nick Schmerr** [11], **Fabian Zander** [6], **Jennifer Inman** [9]

[1]Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland, United States, [2]Department of Electrical and Electronic Engineering, Imperial College London, London, United Kingdom, [3]Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, IRD, Géoazur, Valbonne, France, [4]School of Earth and Planetary Science, Curtin University, Perth, Australia, [5]Los Alamos National Laboratory, New Mexico, United States, [6]Institute for Advanced Engineering and Space Sciences, University of Southern Queensland, Queensland, Australia, [7]United States Department of Defence (KBR Consultant), Las Cruces, New Mexico, United States, [8]NASA Ames Research Center, Moffett Field, California, United States , [9]NASA Langley Research Center, Hampton, Virginia, United States, [10]Raspberry Shake, S.A., Panama, [11]Department of Geology, University of Maryland, College Park, Maryland, United States

Author contributions: *Conceptualization*: Fernando, Lewis, Schmerr. *Methodology*: Fernando, Buttsworth, Hicks, Johnson, McCleary, Zander. *Formal Analysis*: Charalambous, Saliby, Petricca, Larmat, Fernando. *Writing - Original draft*: Fernando, Charalambous.

Abstract Hypersonic re-entries of spacecraft are valuable analogues for the identification and tracking of natural meteoroids re-entering the Earth's atmosphere. We report on the detection of seismic and acoustic signals from the OSIRIS-REx landing sequence, acquired near the point of peak capsule heating and recorded using a fully off-grid Raspberry PiShake sensor. This simple setup is able to record the salient features of both the seismic and acoustic wavefields, including the primary shockwave, later reverberations, and possible locally induced surface waves. Peak overpressures of 0.7 Pa and ground velocities of $2\text{x}10^{-6}$ m/s yield lower bound on the air-to-ground coupling factor between 3 and 44 Hz of $1.4\text{x}10^{-6}$ m/s/Pa, comparable to results from other re-entries.

Production Editor:
Gareth Funning
Handling Editor:
Yen Joe Tan
Copy & Layout Editor:
Oliver Lamb

Signed reviewer(s):
David Fee

Received:
November 15, 2023
Accepted:
February 22, 2024
Published:
March 26, 2024

1 Introduction

1.1 Seismoacoustic measurements of hypervelocity re-entry

Seismic and acoustic measurements are invaluable tools for identifying and locating meteoroids entering the Earth's atmosphere (Edwards et al., 2008). Unlike optical techniques, seismoacoustic techniques allow over-the-horizon measurements to be made, and can continue to track bolides during their dark-flight phase.

The entry, descent, and landing (EDL) of artificial spacecraft can serve as an analogue for these natural meteoroid events, enabling calibration of seismoacoustic measurements using an object of known trajectory, mass, and dimensions (Silber et al., 2023).

However, very rarely do re-entering spacecraft approach velocities representative of naturally-occurring meteorites (>11 km/s, Ceplecha et al., 1998). The exceptions to this are capsules re-entering on interplanetary (as opposed to de-orbital) trajectories.

Such encounters are extremely rare, having occurred only four times on Earth. Sample return capsules from the Genesis (ReVelle et al., 2005), Stardust (ReVelle and Edwards, 2007), Hayabusa (Yamamoto et al., 2011), and Hayabusa2 (Sansom et al., 2022) missions underwent EDLs at around 12 km/s, on the lower end of the velocity distributions of natural meteoroids, though still somewhat representative.

In each case, seismic and acoustic measurements enabled information about the capsule's hypersonic dynamics and the propagation of the sonic boom shockwave to be collected. Ironically, these EDL events are much more common on other planets visited by human spacecraft, but only one has been (unsuccessfully) instrumented (Fernando et al., 2021, 2022).

1.2 The OSIRIS-REx Entry, Descent, and Landing Sequence

The OSIRIS-REx (ORX) sample return capsule was scheduled to re-enter the Earth's atmosphere at 14:41:55 UTC on 2023-09-24, carrying samples from asteroid Bennu (Lauretta et al., 2017). Its atmospheric interface was expected to occur off the coast of San Francisco, California, at an altitude of ∼133 km and a velocity of

*Corresponding author: bfernan9@jh.edu

approximately Mach 25 (43,000 km/h; 11.9 km/s).

ORX's heat shield was expected to experience peak frictional heating from the atmosphere (~3100 K) at a speed of Mach 30 (39,000 km/h; 10.8 km/s) ~62 km above northern Nevada around 14:42:45 UTC, before continuing downrange to a soft landing at the Dugway Proving Ground in Utah at 14:55 UTC (Ajluni et al., 2015).

1.3 Project aims

This project aimed to co-locate a seismoacoustic station with an optical tracking station close to the point of peak heating, in order to study the re-entry process at the point where the maximum amount of kinetic energy is being dissipated into the atmosphere. Exact co-location of acoustic and seismic measurements enables estimation of coupling parameters across the surface interface, helping to constrain how incident acoustic signals produce their seismic counterparts. This is particularly useful when detecting natural meteoroids given that the worldwide seismic network is much denser than its acoustic equivalent.

Whilst other instrumentation campaigns were planned to record seismic and acoustic signatures using more conventional deployments, these were not co-located with an optical tracking station (Silber et al., 2023).

The nearest permanent seismic station was 50 km away (NN.Q11A at Duckwater, Nevada), precluding the use of an existing seismic network to provide local data. Similarly, no permanent infrasound stations were located nearby. Further constraints were imposed on this deployment by the absence of mains power or wired data connections at the optical tracking sites and the stipulation that the data be live-streamed in real-time over the internet for education and outreach purposes.

Our identified solution was to use a low-cost Raspberry PiShake seismic and infrasound sensor[1] coupled to a portable generator and satellite internet connection to fulfil these aims.

In this paper we present the methodology and initial results from this project, whilst also exploring the scalability of a network of this type. For temporary deployments where real-time data access is required (e.g. for monitoring or triggering purposes) in remote areas, such a configuration may serve as a template. This is especially true for phenomena like EDLs where dense instrument spacings are of interest, and the co-location of seismic and acoustic sensors on a single instrument offers both logistical and processing advantages.

This work builds on previous use of distributed off-grid sensors for seismic sensing (e.g. Kong et al. (2016)) and past incorporation of PiShakes into seismic networks (Winter et al., 2021; Mikael, 2020; Manconi et al., 2018; Lecocq et al., 2020). However, it is the first example of which we are aware of a direct PiShake-satellite connection.

[1] Raspberry Shake & Boom, https://manual.raspberryshake.org/boom.html

2 Methodology

2.1 Location

The location of the optical tracking station with which the PiShake (station code: 'RD04A') was co-located was selected by NASA's Scientifically Calibrated In-Flight Imagery (SCIFLI) Team to be close to the point of peak heating in the ORX EDL trajectory, whilst also being remote and far from any artificial light sources.

The selected location was in Eureka County, Nevada (39.264605°N, 116.026934°W), at an elevation of 1843 m AMSL. This site was ~40.5 km laterally offset from the closest point on the nominal EDL trajectory at a bearing of 199° (meaning the minimum source-receiver distance was expected to be 72.2 km). A schematic illustration of the projected EDL trajectory is shown in Fig. 1. Note that the lateral offset was chosen to enable a reduced slewing rate across the sky for the optical tracking instruments.

A range of hills with peaks up to 2900 m (700 m prominence) to the north and north-east at a range of 5-10 km were also noted.

The chosen site was a flat, dry bed, which was identified as having a surface of unconsolidated alluvium. A small section of the ground was artificially brushed clean and smoothed before the seismometer was deployed. P-wave speeds in unsaturated northern Nevada alluvium are reported in the literature as varying between 365 and 1035 m/s (Allander and Berger, 2009).

At a location this distance from and altitude below the EDL track, we anticipated detection both of the direct sonic boom (on the acoustic sensor) and the induced deformation of the ground (on the seismometer). It was also expected from published literature analysing conventional explosive sources that further features might be detected in the seismic coda, for example coupled surface waves (Novoselov et al., 2020; Langston, 2004). Previous work indicates that these observations are site-specific and hence not a given, with a dependence on both local ground properties and current atmospheric conditions (Wills et al., 2022; Chen et al., 2023).

2.2 Weather

The nearest weather station to the seismometer deployment was at Eureka Airport (39.600506°N, 116.006467°W). The distance between the PiShake location and the airport point is 37.4 km, at a bearing of 2.7° from north. A weather measurement was made at the airport at 14:53 UTC, around 11 minutes after the expected overflight of the capsule.

The recorded air temperature was 8.9°C, with a dewpoint of -1.7°C and a resulting relative humidity of 47%. Barometic pressure was 1022.1 hPa, and the windspeed was recorded as 2.1 m/s from an origin bearing 210°. The resulting surface sound speed is calculated to be 337 m/s.

2.3 Setup

A PiShake 'Shake and Boom' equipped with an infrasound sensor and vertical component geophone was

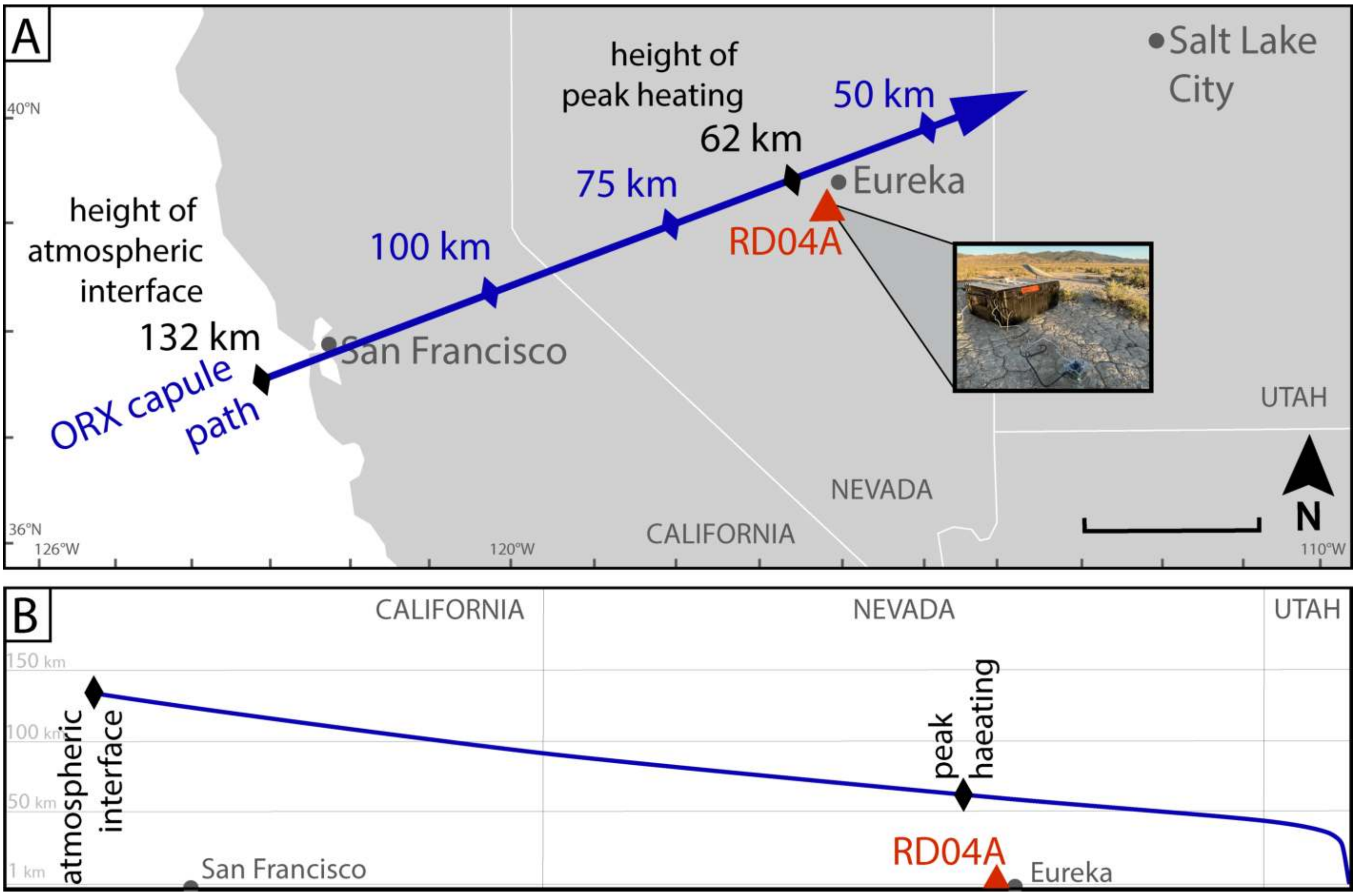


Figure 1 Schematic views of the pre-landing projected ORX EDL path in blue, showing top-down (upper panel) and side-on (lower panel) views. Capsule heights above sea level are indicated along the trajectory. The total length of the path flown after atmospheric interface is approximately 1500 km.

used in this experiment. The instrument was levelled on the ground but not rotationally oriented due to the absence of horizontal component geophones. This setup is shown in Fig. 2. Due to the soil conditions, the sensor could not be feasibly anchored into the ground and simply rested on the surface.

Both the acoustic and seismic sensors sampled using default settings, at 100 samples/second with an estimated bandwidth of -3dB between 0.7 and 44 Hz in velocity for the geophone and -3dB between 1 and 44 Hz for the infrasound sensor.

The sensor was powered by connection to a portable generator approximately 12 m away, which was shielded by makeshift acoustic baffles (bins and camping chairs). A direct connection into a Starlink terminal provided real-time access to the data and livestreaming capability over the PiShake website. Due to the lack of multiple ethernet ports on the terminal, instrument configuration to update metadata could not be done on-site and was executed remotely.

Timing was executed via the instrument's default NTP over the Starlink connection. Data collected between 14:00 and 15:00 UTC is considered reliable, and the background noise levels are representative of the environmental conditions; data are available from before 14:00 UTC on the day of landing but are contaminated by noise from the site setup.

Figure 2 Instrument setup. The PiShake is visible as the transparent box in the foreground, with the geophone and infrasound sensor mounted within the same instrument and levelled using the built-in spirit level. Power (white cable) and data (black cable) are routed through the weatherproof black box. The generator and acoustic baffles are not shown, but the Starlink connection is visible in the background. The unconsolidated alluvial surface and range of hills with peaks 5-10 km away are also visible. Image direction: north (toward EDL trajectory).

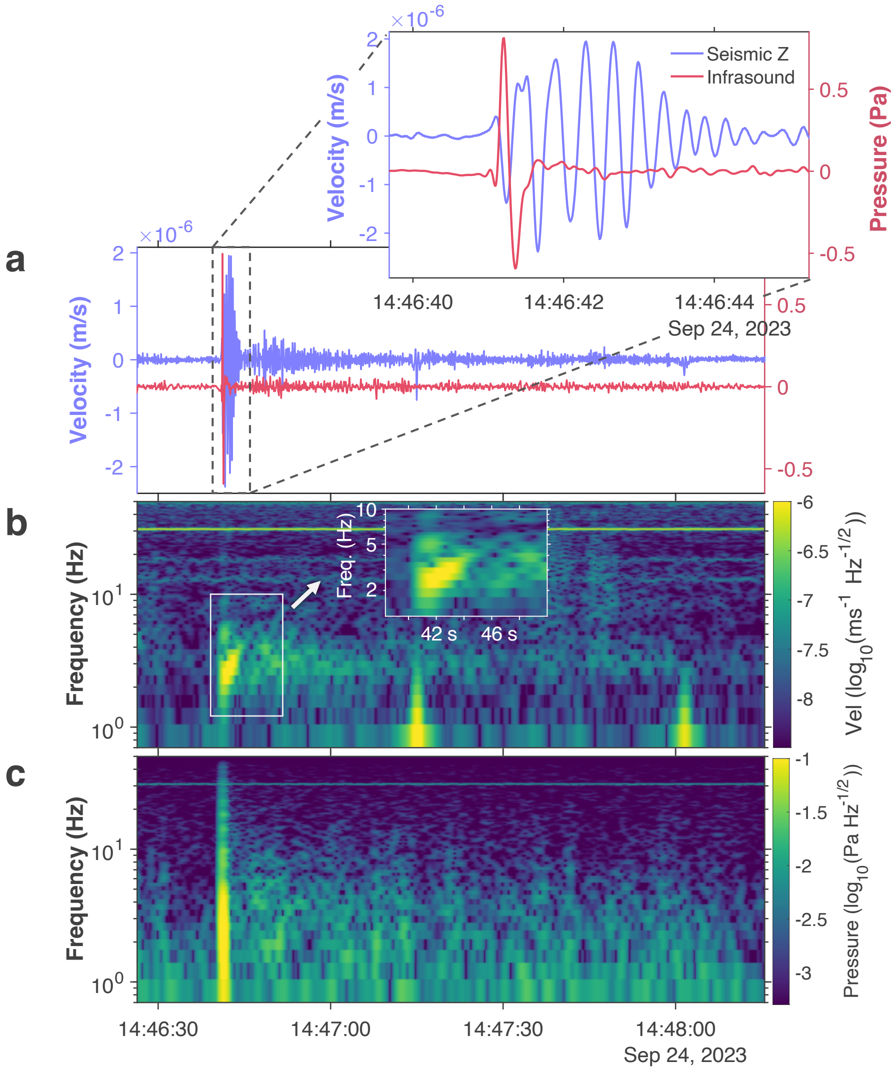


Figure 3 (a) Traces of the seismic (blue) and acoustic (red) data, inset is a detail of the first arrivals. (b) Spectrogram of the seismic data, inset shows a detail of the seismic chirp observed which lasts around four seconds and is dispersive. Bright, vertical spikes at 14:47:15 and 14:48:00 UTC are glitches in the system electronics. Horizontal lines are resonances produced by the generator. (c) Acoustic data, showing a single impulsive arrival at 14:46:45 UTC with no clear coda, but potentially elevated noise levels post-arrival.

3 Results

Pre-landing estimates suggested that any sonic boom would likely arrive at the deployment location around 240 seconds (4 minutes) after the point of peak heating or around 14:46:45 UTC.

Data from the seismic and acoustic instruments are shown in Fig. 3. A sharp peak in both datasets, elevated significantly above the background noise, is recorded at 14:46:41 UTC. This sound was also recorded by the ground team as a loud ‘popping' noise at 14:46:45 UTC ± 00:00:03; extremely short in duration and lacking any discernible internal structure or audible rumbling.

As best as could be determined by the ground team, this boom originated from the north of the observation station (toward the EDL trajectory), but from the direction of the horizon rather than from an elevated angle. We attribute this observation to the fact that the capsule is not behaving as a point source. Rather, it behaves as an elongated cylindrical source producing a conical shock with a hyperbolic footprint on the ground.

We note that the arrival time of the sonic boom was in very close agreement with our pre-landing prediction. However, we expect that this degree of agreement is not particularly consequential (i.e., cannot be used to confirm that the EDL trajectory was totally nominal), as our pre-landing estimate of the boom propagation time was linear and neglected atmospheric refraction of the shockwave.

3.1 Infrasound data

The infrasound signal displays rounded ‘N-wave' behaviour expected of a downward-propagating sonic boom (Plotkin, 2002), with a rapid overpressure (0.7 Pa) pulse and sharp peak followed by an underpressure trough (0.6 Pa) lasting approximately 0.5 s total.

This shape is characteristic of a shockwave which has been distorted by propagation through a turbulent atmosphere (Pierce and Maglieri, 1972). It is very similar to previously recorded signals from hypersonic re-entries (e.g. ReVelle et al., 2005). Hence, we conclude that the main infrasound signal is the direct detection of the sonic boom from the capsule and not an acoustic reflection (echo) or rumbling produced by the incident acoustic wave upon the ground.

The background infrasound noise level appears to be slightly enhanced at low frequencies (<10 Hz) following the arrival as compared to before, though not enormously so (Fig. 3, $\sim$5 s before and after the infrasound arrival). This feature may be the signature of a low-frequency sub-audible infrasonic rumble or may simply be associated with elevated wind noise.

3.2 Seismic data

The seismic dataset appears to be considerably richer than the infrasound signature, with a signal lasting approximately 120 seconds (2 minutes). The first seismic arrival is exactly coincident with the infrasound arrival. This likely represents the shaking of the surface induced by the overpressure (Ben-Menahem and Singh, 1981).

This conclusion is also supported by the polarity of the signal, with the downward motion at first arrival corresponding to the displacement downward of the sensor and ground in response to the acoustic overpressure. Peak shaking of 2x10^{-6} m/s is observed, with the boom itself lasting approximately 5 seconds.

The complex coda is likely to have multiple origins, including the excitation of surface waves (Cook et al., 1972), scattering of the shockwave in the atmosphere (Garcia et al., 2022), reflections of the direct shockwave off of topography (Emmanuelli et al., 2021) and of the transmitted shockwave off sub-surface geological features, and the gradual restoration of the equilibrium surface position following compliance-induced deformation.

A short, dispersive, chirp-like signal is apparent between approximately 1 and 7 Hz in the 3-4 seconds following the initial seismic arrival (see inset spectrogram, beginning at 14:46:41 UTC), with higher frequencies arriving later than lower ones. Two potential chirping structures are also observed later in the data (around 14:46:46 and 14:46:50 UTC), though at much lower signal-to-noise ratios.

'Normal' dispersion in chirp structures is indicative of higher frequencies propagating more slowly than lower ones. Such behaviour would be expected for seismic surface waves propagating in the uppermost layers of the ground, where the gradient in sound speed with depth is significant on the scale of the seismic wavelength. Because these surface phases do not arrive before the initial boom, we conclude that they are produced locally to the receiver; as seismoacoustic coupling in the far-field directly below the ORX trajectory would likely see coupled surface waves 'overtaking' the boom due to the higher propagation speeds in the ground.

Similar features, identified as Airy waves, are seen during the Stardust EDL by Edwards et al. (2007), whilst Novoselov et al. (2020) identify Stoneley waves in seismic coda generated by seismoacoustic coupling. These propagate in the thin, low-velocity surface layers where the shear velocity approaches the acoustic wavespeed in air (Wills et al., 2022). This is comparable to the geological setting here, with the PiShake sensor resting on a low-velocity alluvial layer.

3.3 Air-to-ground coupling

The co-location of seismic and acoustic instruments allows us to estimate the ground compliance (the ground's response to the pressure loading from the shockwave at the surface (Sorrells, 1971; Kenda et al., 2020). A number of physical phenomena contribute to the compliance, here we consider the inertial effects originating from the continuity of normal stress and displacement at the ground surface. We note that measurements of seismoacoustic coupling strength are in general very sensitive, in particular to the frequency bands considered, surface topography, and wavefront shape/incident angle (Matoza and Fee, 2014; Bishop et al., 2022).

We estimate the inertial effects by considering how the vertical deformation recorded by the seismometer is related to the surface overpressure. The ground is modelled as a homogeneous, isotropic half-space and the shockwave is modelled as a planar wave. Following Kenda et al. (2020), the compliance K_v is then:

$$K_v = 2c\frac{1-\nu^2}{E}, \tag{1}$$

where E is the Young's modulus, ν is the Poisson ratio, and c is the advection speed of the pressure loading.

We choose elastic properties for the subsurface corresponding to canonical values for soft superficial alluvium, with V_p = 585 m/s and V_s = 350 m/s, ν = 0.22, and E = 0.4488 MPa. These are commensurate with detailed surveys from the wider region (Allander and Berger, 2009). Local variations in these values may be substantial, but would require a full geophysical survey to better constrain.

These values lead to a compliance dependent on the advection speed c of:

$$K_v = 4.23\text{x}10^{-9}c \ \text{ m/s/Pa} \tag{2}$$

As we expect the shockwave to propagate at or faster than the speed of sound, we use a value for the advective speed c of 337 m/s, as derived in Sec. 2.2. This is very much a lower limit, as the actual speed of the shockwave is not simple to measure. We also note that this value is much faster than conventional derivations of the compliance, which use wind-driven ground deformation as a source ($c \sim$2 m/s in Kenda et al., 2020).

These parameters yield a minimum value for the ground compliance of 1.4x10^{-6}. This gives us a calculated minimum value for the air-to-ground coupling of 4x10^{-6} m/s/Pa, using the vertical component of the velocity only.

This value is comparable to that derived from the Stardust EDL by Edwards et al. (2007) of 7.3$\pm$ 0.2x10^{-6} m/s/Pa, which used a similar capsule and trajectory but required a far more complex deployment to estimate.

3.4 Seismoacoustic noise

For completeness, we comment on a number of seismoacoustic noise sources which are apparent in the wider dataset.

Strong resonances in both instruments are observed at 30 Hz, with weaker resonances in the seismic data at ∼11 Hz and ∼19 Hz. These appear as horizontal lines in Fig. 3. These are identified as coming from the generator, and over a longer timescale (not shown here), they display subtle changes in frequency as the generator load varies with changes in the power demand of the optical tracking instruments.

Occasional spikes in the seismic power are also observed, for example at 14:46:45 UTC. These are also thought to be electromagnetic glitches associated with rapid changes in the generator's load. These appear as vertical spikes lines in Fig. 3. One glitch at 14:47:15 UTC partially overprints the seismic coda, though well beyond the point at which the identified surface wave phases have dropped below the noise floor.

4 Discussion

4.1 Scientific utility

This deployment demonstrated the ability of a fully off-grid Raspberry PiShake 'Shake and Boom' sensor to capture valuable data from a transient seismoacoustic event, whilst also making said data publicly accessible via livestream.

Whilst naturally limited in sensitivities to long periods (lower than 1 Hz), this work also demonstrates the ability of the PiShake instrument to capture many of the notable features in the wavetrain, from the initial rounded N-wave to the coda likely associated with Stoneley waves propagating in the low-velocity subsurface. Whilst our single station does not offer the same seismic insight as arrays or co-located broader spectrum instruments would, these features of the wavetrain are recorded comparably to past studies (e.g. ReVelle et al., 2005; ReVelle and Edwards, 2007).

These include the primary shockwave (0.7 Pa overpressure and 2x10^{-6} m/s peak ground velocity), an extended seismic coda, and possible air-to-ground surface wave phases. We derive a lower bound on the air-to-ground coupling ratio of 4x10^{-6} m/s/Pa, comparable to previous capsule re-entries.

4.2 Deployment suggestions

The use of a surface instrument which is locally powered and not hard-coupled to the ground obviously brings with it disadvantages, not least of which is the generator noise apparent on both sensors. For those looking to undertake similar deployments, we also make the following suggestions:

- The addition of a wind cover to the instrument would likely substantially reduce the noise levels of both the acoustic and seismic data.

- Better anchoring of the instrument into the ground would be expected to especially benefit the seismic data.

- The use of a well-tuned and lubricated generator can minimise the amount of acoustic noise produced, enabling data to be recorded with higher fidelity. A rechargable battery or solar panel would of course decrease noise levels even further.

- A dedicated instrument power supply (whether battery or generator) can avoid fluctuations in the generator load which lead to variations in the generator resonant frequencies. Such variations make removal of the generator noise harder.

- For standard Starlink terminals, the single ethernet output port means that LAN configuration of the instrument must be modified, i.e. it is not possible to connect a laptop and a PiShake to the terminal simultaneously without additional hardware. This can be avoided by connecting the PiShake through a laptop to enable metadata edits, or configuring location and elevation parameters on a different network prior to deployment.

4.3 Scaling to larger arrays

The nominal data uplink rate from a Raspberry Pi Shake&Boom sensor of this type is approximately 2.8 kb/s. A standard Starlink connection comes with a minimum expected data rate of 5 Mb/s. As such, data volumes are unlikely to prove problematic for realistic array sizes (<1000 instruments). We again note, however, that additional hardware would be required to enable multiple wired uploads through a Starlink connector due to the single ethernet port on the terminal.

The power requirements of the sensor are at least 5.0 V DC at 2.5 A, for an electrical power of 12.5W. As such, an intermediate-sized array (∼100 instruments) could likely be powered from a single generator for a number of hours. For signals such as those discussed in this paper, that is likely to be more than sufficient.

Finally, we suggest the co-location of 3D PiShakes with PiShake infrasound sensors would also be advantageous, enabling the the air-to-ground coupling factor to be more robustly estimated as all three components of ground displacement can be considered. A larger array would also enable a more realistic estimation of the shock velocity.

5 Conclusions

We have demonstrated that a fully off-grid, livestreaming Raspberry PiShake sensor was able to capture the notable seismic and acoustic features within the waveforms produced by the OSIRIS-REx EDL sequence. A classical rounded 'N-wave' was recorded by the acoustic sensor (peak overpressure 0.7 Pa), and the reverse polarity signal was recorded by the seismometer (peak downward ground velocity 2x10^{-6} m/s), indicating the recording of a sonic boom impacting the ground with an acoustic coupling of at least 4x10^{-6} m/s/Pa.

No unambiguous acoustic coda is recorded, though a possible detection of low-frequency rumbling is made. An extended seismic coda of more complex origin is also recorded, and lasts several minutes.

Although the initial configuration required substantial bespoke setup with satellite internet and data connections, the marginal cost or challenge of adding extra instruments was small. As such, we believe there is substantial potential for low-cost arrays of this type to be scaled to larger sizes when there is a scientific need to record seismoacoustic phenomena at high SNR in remote locations. Such potential has already been demonstrated with conventional seismometers (Busby and Aderhold, 2020) but not to our knowledge with PiShake-type arrays.

6 Environmental impact

As part of efforts to make scientific research more environmentally accountable and sustainable, we have estimated the 'carbon cost' (expressed as tCO_2e) directly associated with this paper. As is common with such projects, scope definition is challenging. Therefore, we focus on carbon costs directly attributed to this project (i.e., those which would not have otherwise been incurred). We exclude background costs such as the instrument manufacture.

We estimate the total equivalent CO_2 burden of this work at approximately 1.0 tCO_2e, made up of:

- 0.5 tCO_2e associated with a round-trip flight from Baltimore, Maryland to Salt Lake City, Utah; in economy class as estimated by the IATA CO_2 connect calculator [2]
- 0.5 tCO_2e associated with driving an SUV off-road field vehicle from Salt Lake City, Utah to the field site near Eureka, Nevada (approximately 800 miles total); estimated using the US EPA Equivalencies Calculator [3]. Note that this ride was shared by several other instrument groups, we quote here the total cost.
- A negligible amount (<0.01 tCO_2e) produced by running the gasoline-powered generator for one hour.

The above values are estimates only, and we note that differences between calculation methodologies can be substantial.

Acknowledgements

The authors are deeply grateful to the townspeople of Eureka, Nevada, for providing a welcome and hospitable fieldwork environment, and for lending us a shovel!

We are also grateful to members of NASA Langley's SCIFLI team for their assistance in experiment design and information on the ORX EDL trajectory; as well as David Fee and another anonymous reviewer.

BF is funded by the Blaustein Fellowship Program in the Department of Earth and Planetary Sciences and Johns Hopkins University. CS acknowledges the support of ANR (MAGIS, ANR-19-CE31-0008-08).

[2]https://www.iata.org/en/services/statistics/intelligence/co2-connect/
[3]https://www.epa.gov/energy/greenhouse-gas-equivalencies-calculator

Data and code availability

PiShake data are available freely online from the Raspberry Shake Data Centre and may be accessed via the FDSN (http://www.fdsn.org/datacenters/detail/RASPISHAKE/).

The relevant station code is RD04A with channels EHZ (vertical-component geophone) and HDF (infrasound). Note that the current station location may display as Baltimore, Maryland, where this instrument lives on a more permanent basis. Response removal and other useful tools can be found on the PiShake website at https://manual.raspberryshake.org/metadata.html#example-obspy-code-for-removing-response.

Competing interests

The authors declare no competing interests.

References

Ajluni, T., Everett, D., Linn, T., Mink, R., Willcockson, W., and Wood, J. OSIRIS-REx, returning the asteroid sample. In *2015 IEEE aerospace conference*, pages 1–15. IEEE, 2015. doi: 10.1109/AERO.2015.7118988.

Allander, K. K. and Berger, D. L. Seismic velocities and thicknesses of alluvial deposits along Baker Creek in the Great Basin National Park, East-Central Nevada. *Technical Report, U. S. Geological Survey*, 2009. doi: https://doi.org/10.3133/ofr20091174.

Ben-Menahem, A. and Singh, S. J. *Seismic waves and sources*. Springer Science & Business Media, 1981. doi: https://doi.org/10.1007/978-1-4612-5856-8.

Bishop, J. W., Fee, D., Modrak, R., Tape, C., and Kim, K. Spectral element modeling of acoustic to seismic coupling over topography. *Journal of Geophysical Research: Solid Earth*, 127(1):e2021JB023142, 2022. doi: https://doi.org/10.1029/2021JB023142.

Busby, R. W. and Aderhold, K. The Alaska transportable array: As built. *Seismological Research Letters*, 91(6):3017–3027, 2020. doi: https://doi.org/10.1785/0220200154.

Ceplecha, Z., Borovička, J., Elford, W. G., ReVelle, D. O., Hawkes, R. L., Porubčan, V., and Šimek, M. Meteor phenomena and bodies. *Space Science Reviews*, 84:327–471, 1998. doi: https://doi.org/10.1023/A:1005069928850.

Chen, T., Larmat, C., Blom, P., and Zeiler, C. Seismoacoustic Analysis of the Large Surface Explosion Coupling Experiment Using a Large-N Seismic Array. *Bulletin of the Seismological Society of America*, 113(4):1692–1701, 2023. doi: https://doi.org/10.1785/0120220262.

Cook, J., Goforth, T., and Cook, R. Seismic and underwater responses to sonic boom. *The Journal of the Acoustical Society of America*, 51(2C):729–741, 1972. doi: https://doi.org/10.1121/1.1912906.

Edwards, W. N., Eaton, D. W., McCausland, P. J., ReVelle, D. O., and Brown, P. G. Calibrating infrasonic to seismic coupling using the Stardust sample return capsule shockwave: Implications for seismic observations of meteors. *Journal of Geophysical Research*, 112(B10), 2007. doi: 10.1029/2006jb004621.

Edwards, W. N., Eaton, D. W., and Brown, P. G. Seismic observations of meteors: Coupling theory and observations. *Reviews of Geophysics*, 46(4), 2008. doi: https://doi.org/10.1029/2007RG000253.

Emmanuelli, A., Dragna, D., Ollivier, S., and Blanc-Benon, P. Characterization of topographic effects on sonic boom reflection by resolution of the Euler equations. *The Journal of the Acoustical Society of America*, 149(4):2437–2450, 2021. doi: https://doi.org/10.1121/10.0003816.

Fernando, B., Wójcicka, N., Froment, M., Maguire, R., Stähler, S. C., Rolland, L., Collins, G. S., Karatekin, O., Larmat, C., Sansom, E. K., et al. Listening for the landing: Seismic detections of Perseverance's arrival at Mars with InSight. *Earth and Space Science*, 8(4):e2020EA001585, 2021. doi: https://doi.org/10.1029/2020EA001585.

Fernando, B., Wójcicka, N., Maguire, R., Stähler, S. C., Stott, A. E., Ceylan, S., Charalambous, C., Clinton, J., Collins, G. S., Dahmen, N., et al. Seismic constraints from a Mars impact experiment using InSight and Perseverance. *Nature Astronomy*, 6(1):59–64, 2022. doi: https://doi.org/10.1038/s41550-021-01502-0.

Garcia, R. F., Daubar, I. J., Beucler, É., Posiolova, L. V., Collins, G. S., Lognonné, P., Rolland, L., Xu, Z., Wójcicka, N., Spiga, A., et al. Newly formed craters on Mars located using seismic and acoustic wave data from InSight. *Nature Geoscience*, 15(10):774–780, 2022. doi: https://doi.org/10.1038/s41561-022-01014-0.

Kenda, B., Drilleau, M., Garcia, R. F., Kawamura, T., Murdoch, N., Compaire, N., Lognonné, P., Spiga, A., Widmer-Schnidrig, R., Delage, P., Ansan, V., Vrettos, C., Rodriguez, S., Banerdt, W. B., Banfield, D., Antonangeli, D., Christensen, U., Mimoun, D., Mocquet, A., and Spohn, T. Subsurface Structure at the InSight Landing Site From Compliance Measurements by Seismic and Meteorological Experiments. *Journal of Geophysical Research: Planets*, 125(6), 2020. doi: 10.1029/2020je006387.

Kong, Q., Allen, R. M., Schreier, L., and Kwon, Y.-W. MyShake: A smartphone seismic network for earthquake early warning and beyond. *Science advances*, 2(2):e1501055, 2016. doi: https://doi.org/10.1126/sciadv.1501055.

Langston, C. A. Seismic ground motions from a bolide shock wave. *Journal of Geophysical Research: Solid Earth*, 109(B12), 2004. doi: https://doi.org/10.1029/2004JB003167.

Lauretta, D., Balram-Knutson, S., Beshore, E., Boynton, W., Drouet d'Aubigny, C., DellaGiustina, D., Enos, H., Golish, D., Hergenrother, C., Howell, E., et al. OSIRIS-REx: sample return from asteroid (101955) Bennu. *Space Science Reviews*, 212:925–984, 2017. doi: https://doi.org/10.1007/s11214-017-0405-1.

Lecocq, T., Hicks, S. P., Van Noten, K., Van Wijk, K., Koelemeijer, P., De Plaen, R. S., Massin, F., Hillers, G., Anthony, R. E., Apoloner, M.-T., et al. Global quieting of high-frequency seismic noise due to COVID-19 pandemic lockdown measures. *Science*, 369(6509):1338–1343, 2020. doi: https://doi.org/10.1126/science.abd2438.

Manconi, A., Coviello, V., Galletti, M., and Seifert, R. Monitoring rockfalls with the Raspberry Shake. *Earth Surface Dynamics*, 6 (4):1219–1227, 2018. doi: https://doi.org/10.5194/esurf-6-1219-2018.

Matoza, R. S. and Fee, D. Infrasonic component of volcano-seismic eruption tremor. *Geophysical Research Letters*, 41(6): 1964–1970, 2014. doi: https://doi.org/10.1002/2014GL059301.

Mikael, S. Establishment of the Ethiopian Seismic Monitoring Network. *Engineering Archive*, 2020. doi: https://doi.org/10.31224/osf.io/c6y8p.

Novoselov, A., Fuchs, F., and Bokelmann, G. Acoustic-to-seismic ground coupling: coupling efficiency and inferring near-surface properties. *Geophysical Journal International*, 223(1):144–160, 2020. doi: https://doi.org/10.1093/gji/ggaa304.

Pierce, A. D. and Maglieri, D. J. Effects of atmospheric irregularities on sonic-boom propagation. *The Journal of the Acoustical Society of America*, 51(2C):702–721, 1972. doi: https://doi.org/10.1121/1.1912904.

Plotkin, K. J. State of the art of sonic boom modeling. *The Journal of the Acoustical Society of America*, 111(1):530–536, 2002. doi: https://doi.org/10.1121/1.1379075.

ReVelle, D. and Edwards, W. Stardust—An artificial, low-velocity "meteor" fall and recovery: 15 January 2006. *Meteoritics & Planetary Science*, 42(2):271–299, 2007. doi: https://doi.org/10.1111/j.1945-5100.2007.tb00232.x.

ReVelle, D., Edwards, W., and Sandoval, T. Genesis—An artificial, low velocity "meteor" fall and recovery: September 8, 2004. *Meteoritics & Planetary Science*, 40(6):895–916, 2005. doi: https://doi.org/10.1111/j.1945-5100.2005.tb00162.x.

Sansom, E. K., Devillepoix, H. A., Yamamoto, M.-y., Abe, S., Nozawa, S., Towner, M. C., Cupák, M., Hiramatsu, Y., Kawamura, T., Fujita, K., et al. The scientific observation campaign of the Hayabusa-2 capsule re-entry. *Publications of the Astronomical Society of Japan*, 74(1):50–63, 2022. doi: https://doi.org/10.1093/pasj/psab109.

Silber, E. A., Bowman, D. C., and Albert, S. A Review of Infrasound and Seismic Observations of Sample Return Capsules Since the End of the Apollo Era in Anticipation of the OSIRIS-REx Arrival. *Atmosphere*, 14(10):1473, 2023. doi: https://doi.org/10.3390/atmos14101473.

Sorrells, G. G. A Preliminary Investigation into the Relationship between Long-Period Seismic Noise and Local Fluctuations in the Atmospheric Pressure Field. *Geophysical Journal of the Royal Astronomical Society*, 26(1-4):71–82, 1971. doi: 10.1111/j.1365-246X.1971.tb03383.x.

Wills, G., Nippress, A., Green, D. N., and Spence, P. J. Site-specific variations in air-to-ground coupled seismic arrivals from the 2012 October 16 explosion at Camp Minden, Louisiana, United States. *Geophysical Journal International*, 231(1):243–255, 2022. doi: https://doi.org/10.1093/gji/ggac184.

Winter, K., Lombardi, D., Diaz-Moreno, A., and Bainbridge, R. Monitoring icequakes in East Antarctica with the raspberry shake. *Seismological Research Letters*, 92(5):2736–2747, 2021. doi: https://doi.org/10.1785/0220200483.

Yamamoto, M.-y., Ishihara, Y., Hiramatsu, Y., Kitamura, K., Ueda, M., Shiba, Y., Furumoto, M., and Fujita, K. Detection of acoustic/infrasonic/seismic waves generated by hypersonic re-entry of the HAYABUSA capsule and fragmented parts of the spacecraft. *Publications of the Astronomical Society of Japan*, 63(5): 971–978, 2011. doi: https://doi.org/10.1093/pasj/63.5.971.

REPORT
doi:10.26443/seismica.v3i1.1143

VIP - Variational Inversion Package with example implementations of Bayesian tomographic imaging

Xin Zhang * [1,2], **Andrew Curtis** [2]

[1]School of Engineering and Technology, China University of Geosciences, Beijing, China, [2]School of GeoSciences, University of Edinburgh, UK

Author contributions: *Conceptualization*: Xin Zhang, Andrew Curtis. *Writing - Original draft*: Xin Zhang.

Abstract Bayesian inference has become an important methodology to solve inverse problems and to quantify uncertainties in their solutions. Variational inference is a method that provides probabilistic, Bayesian solutions efficiently by using optimisation. In this study we present a Python Variational Inversion Package (VIP), to solve inverse problems using variational inference methods. The package includes automatic differential variational inference (ADVI), Stein variational gradient descent (SVGD) and stochastic SVGD (sSVGD), and provides implementations of 2D travel time tomography and 2D full waveform inversion including test examples and solutions. Users can solve their own problems by supplying an appropriate forward function and a gradient calculation code. In addition, the package provides a scalable implementation which can be deployed easily on a desktop machine or using modern high performance computational facilities. The examples demonstrate that VIP is an efficient, scalable, extensible and user-friendly package, and can be used to solve a wide range of low or high dimensional inverse problems in practice.

Production Editor:
Gareth Funning
Handling Editor:
Paula Koelemeijer
Copy & Layout Editor:
Hannah F. Mark

Received:
October 27, 2023
Accepted:
May 2, 2024
Published:
May 29, 2024

1 Introduction

In a variety of academic and practical applications that concern the Earth's subsurface we wish to find answers to specific scientific questions. In the geosciences this is often achieved by imaging subsurface properties using data recorded on the surface, and by interpreting those images to address questions of interest. The subsurface is usually parameterised in some way, and a physical relationship is defined that predicts data that would be recorded for any particular set of model parameters, while the inverse relationship can not be determined uniquely. Once real data have been observed, the imaging problem is thus established as an inverse problem (Tarantola, 2005).

Because of non-linearity in the physical relationship, insufficient data coverage and noise in the data, inverse problems almost always have non-unique solutions: many sets of parameter values can fit the data to within their uncertainty. It is therefore important to characterize the family of possible solutions (in other words, the solution uncertainty) in order to interpret the results with the correct level of confidence, and to provide well-justified and robust answers to the scientific questions (Arnold and Curtis, 2018).

Solutions to an inverse problem are often found by seeking an optimal set of parameter values that minimizes the difference or misfit between observed data and model-predicted data to within the data noise. Since most inverse problems have non-unique solutions, some form of regularization is often imposed on the parameters in order to make the computational solution unique (Aki and Lee, 1976; Tarantola, 2005; Aster et al., 2018). Many codes have been developed using this class of methods (Rawlinson, 2005; Rücker et al., 2017; Afanasiev et al., 2019; Wathelet et al., 2020; Komatitsch et al., 2023). However, since regularization is often chosen using ad-hoc criteria, these methods produce deliberately biased results, and valuable information can be concealed in the process (Zhdanov, 2002). Moreover, no such optimisation method can provide accurate estimates of uncertainty. To overcome these issues, the SOLA-Backus-Gilbert inversion method has recently been applied to large scale linearised tomographic problems. This method evaluates the weighted average of the true model parameters and provides both resolution and uncertainty estimates (Zaroli, 2016; Zaroli et al., 2017). In addition, the method does not require regularization and can be conducted in a parameter-free way which avoids bias caused by parameterisation (Zaroli, 2019). Unfortunately, the method is only developed for linear problems; since most Geophysical problems are significantly nonlinear, our goal is to provide methods that estimate solutions and uncertainties for that class of problems.

Bayesian inference solves both linear and nonlinear inverse problems by updating a *prior* probability density function (pdf) with new information contained in the data to produce a *posterior* pdf which describes the full state of information about the parameters post inversion (Tarantola, 2005). If we define the prior pdf as $p(\mathbf{m})$, the posterior pdf $p(\mathbf{m}|\mathbf{d}_{\mathrm{obs}})$ can be computed using Bayes' theorem:

$$p(\mathbf{m}|\mathbf{d}_{\mathrm{obs}}) = \frac{p(\mathbf{d}_{\mathrm{obs}}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d}_{\mathrm{obs}})} \tag{1}$$

*Corresponding author: xzhang@cugb.edu.cn

where $p(\mathbf{d}_{\text{obs}}|\mathbf{m})$ is the *likelihood* function which describes the probability of observing the recorded data $\mathbf{d}_{\text{obs}}$ if model parameters took the values in $\mathbf{m}$, and $p(\mathbf{d}_{\text{obs}})$ is a normalization factor called the *evidence*. This posterior pdf describes the full uncertainty in parameter values by combining the prior information and the uncertainty contained in the data.

Markov chain Monte Carlo (McMC) is one commonly-used method to solve Bayesian inference problems and has been used widely in many fields. The method constructs a set (chain) of successive samples that are distributed according to the posterior pdf by performing a structured random walk through parameter space (Brooks et al., 2011); thereafter, these samples can be used to estimate statistical information about parameters in the posterior pdf (Mosegaard and Tarantola, 1995; Tarantola, 2005) and to find answers to specific scientific questions (Arnold and Curtis, 2018; Siahkoohi et al., 2022b; Zhang and Curtis, 2022; Zhao et al., 2022b; McKean et al., 2023). The Metropolis-Hastings algorithm is one such method that originates from physics (Metropolis and Ulam, 1949; Hastings, 1970), and has been applied to a range of geophysical inverse problems (Mosegaard and Tarantola, 1995; Malinverno et al., 2000; Andersen et al., 2001; Mosegaard and Sambridge, 2002; Sambridge and Mosegaard, 2002; Ramirez et al., 2005; Gallagher et al., 2009). However, the algorithm becomes inefficient in high dimensional space because of poor scaling due to its random walk behaviour.

In order to solve Bayesian inference problems more efficiently, a variety of more advanced methods have been introduced to geophysics, such as reversible-jump McMC (Green, 1995; Malinverno, 2002; Bodin and Sambridge, 2009; Galetti et al., 2015; Zhang et al., 2018b), Hamiltonian Monte Carlo (Duane et al., 1987; Sen and Biswas, 2017; Fichtner et al., 2018; Gebraad et al., 2020), Langevin Monte Carlo (Roberts et al., 1996; Siahkoohi et al., 2020), stochastic Newton McMC (Martin et al., 2012; Zhao and Sen, 2019), and parallel tempering (Hukushima and Nemoto, 1996; Dosso et al., 2012; Sambridge, 2013). Gaussian process models have also been used to solve linearised probabilistic problems (Valentine and Sambridge, 2020). Based on these studies a range of methods and codes have been developed to solve geophysical inverse problems using McMC (Bodin and Sambridge, 2009; Shen et al., 2012; Hawkins and Sambridge, 2015; Zhang et al., 2018b; Zunino et al., 2023). Nevertheless, these papers mainly address 1D, 2D or sparsely-parametrised 3D spatial imaging problems; Bayesian solutions to large scale problems (e.g., those involving thousands of parameters to be estimated) remain intractable because of their unaffordable computational cost due to the curse of dimensionality (Curtis and Lomax, 2001).

In an attempt to improve the efficiency of Bayesian inference for certain types of problems, variational inference has been introduced to geophysics as an alternative to McMC. In variational inference one seeks a best approximation to the posterior pdf within a predefined family of (simplified) probability distributions by minimizing the difference between the approximating pdf and the posterior pdf (Bishop, 2006; Blei et al., 2017). One commonly-used measure of the difference between the pdfs is the Kullback-Leibler (KL) divergence (Kullback and Leibler, 1951) as it is easier to estimate computationally than other measures. Variational inference therefore solves Bayesian inference problems by minimizing the KL divergence, which is an optimisation rather than a stochastic sampling problem. The method has been demonstrated to be computationally more efficient and more scalable to high dimensionality in some classes of problems (Bishop, 2006; Zhang et al., 2018a). The method can also be applied to large datasets by dividing the data set into random minibatches and using stochastic and distributed optimisation (Robbins and Monro, 1951; Kubrusly and Gravier, 1973). By contrast, the same strategy cannot easily be used for McMC because it breaks the detailed balance condition required by most McMC methods (O'Hagan and Forster, 2004). In addition, variational inference methods can usually be parallelized at the individual sample level, whereas in McMC this cannot be achieved because of dependence between successive samples.

Variational inference has been applied to a range of geophysical inverse problems. Nawaz and Curtis (2018) used *mean-field* variational inference to invert for subsurface geological facies distributions and petrophysical properties using seismic data, with further developments by Nawaz and Curtis (2019) and Nawaz et al. (2020). Although these methods are computationally efficient, the mean-field approximation ignores correlations between parameters, and the methods of Nawaz and Curtis involved the development of bespoke mathematical derivations and implementations for each class of problem. While these developments result in exceptional speed of calculation, this approach restricts the method to a small range of problems for which correlations are not important and the derivations can be performed (Parisi, 1988; Bishop, 2006; Blei et al., 2017). To extend variational inference to general inverse problems, Kucukelbir et al. (2017) used a Gaussian family in variational inference to create a method called automatic differential variational inference (ADVI), which has been applied to travel time tomography (Zhang and Curtis, 2020a) and earthquake slip inversion (Zhang and Chen, 2022), and extended to the family of sums (mixtures) of multiple Gaussians by Zhao and Curtis (2024). By using a sequence of invertible and differential transforms (called normalizing flows), Rezende and Mohamed (2015) proposed normalizing flow variational inference in which flows (functions, or simply, relationships) are designed which convert a simple initial distribution to an arbitrarily complex distribution that approximates the posterior pdf. In geophysics and related fields the method has been applied to travel time tomography (Zhao et al., 2022a), seismic imaging (Siahkoohi et al., 2020, 2022a), seismic data interpolation (Kumar et al., 2021), transcranial ultrasound tomography (Orozco et al., 2023) and cascading hazards estimation (Li et al., 2023).

By using a set of samples of parameter values (called particles) to represent the density of an approximating pdf, Liu and Wang (2016) introduced a method called Stein variational gradient descent (SVGD), which iteratively updates those particles by minimizing the KL divergence so that the final particle density provides an approximation to the posterior pdf. SVGD has been demonstrated to be an efficient method in a range of geophysical applications, such as travel time tomography (Zhang and Curtis, 2020a), full waveform inversion (FWI) (Zhang and Curtis, 2020b, 2021; Lomas et al., 2023; Wang et al., 2023), earthquake source inversion (Smith et al., 2022), hydrogeological inversion (Ramgraber et al., 2021), post-stack seismic inversion (Izzatullah et al., 2023) and neural network based seismic tomography (Agata et al., 2023). However the method becomes inefficient and inaccurate in high dimensional problems because of the finite number of particles and the practical limitation of computational cost (Ba et al., 2022). To reduce this issue, Gallego and Insua (2018) introduced the stochastic SVGD (sSVGD) method by combining SVGD and McMC: the efficiency of this method has recently been demonstrated when it was used to estimate the first Bayesian solution for a fully nonlinear, 3D FWI problem (Zhang et al., 2023).

Despite these theoretical and practical advances, variational inference has not been widely used in geophysics. This is partly because the method is not easily accessible to non-specialists, and also because there is no common code framework to perform geophysical inversions using the method. In this study we therefore present a Python variational inversion package (VIP), which includes ADVI, SVGD and sSVGD, to make it more straightforward to solve geophysical inverse problems using variational inference methods. The package provides complete implementations of 2D travel time tomography and 2D full waveform inversion problems, including test results for users to check that their implementation is correct. Users can also solve other inverse problems by supplying their own forward functions and gradient calculation codes. In addition, to solve large inverse problems the package is designed in a scalable way such that it can be deployed on a desktop computer as well as in modern high performance computational (HPC) facilities.

In the following section we describe the concept of variational inference, and algorithmic details of ADVI, SVGD and sSVGD. In section 3 we provide an overview of the VIP package, and in section 4 we demonstrate VIP using examples of 2D travel time tomography and 2D full waveform inversion. We thus show that VIP is an efficient, scalable, extensible and user-friendly package that will enable users to solve geophysical inverse problems using variational methods. Making these methods more tractable for practitioners should allow them to be tested on a wide range of problems.

2 Theoretical background

2.1 Variational inference

Variational inference solves Bayesian inference problems using optimisation. To achieve this, we first define a simplified family of pdf's $Q = \{q(\mathbf{m})\}$, for example, the family of all Gaussian distributions. The method then seeks an optimal approximation $q^*(\mathbf{m})$ to the posterior probability distribution $p(\mathbf{m}|\mathbf{d}_{\text{obs}})$ within this family by minimizing the KL divergence between $q(\mathbf{m})$ and $p(\mathbf{m}|\mathbf{d}_{\text{obs}})$:

$$q^*(\mathbf{m}) = \arg\min_{q\in Q} \mathrm{KL}[q(\mathbf{m})||p(\mathbf{m}|\mathbf{d}_{\text{obs}})] \tag{2}$$

The KL divergence measures the difference between two probability distributions:

$$\begin{aligned}\mathrm{KL}[q(\mathbf{m})||p(\mathbf{m}|\mathbf{d}_{\text{obs}})] &= \mathrm{E}_q[\log q(\mathbf{m})] - \mathrm{E}_q[\log p(\mathbf{m}|\mathbf{d}_{\text{obs}})] \\ &= \mathrm{E}_q[\log q(\mathbf{m})] - \mathrm{E}_q[\log p(\mathbf{m},\mathbf{d}_{\text{obs}})] \\ &+ \log p(\mathbf{d}_{\text{obs}})\end{aligned} \tag{3}$$

where $\log p(\mathbf{m},\mathbf{d}_{\text{obs}})$ is the joint distribution of model $\mathbf{m}$ and data $\mathbf{d}_{\text{obs}}$. The expectations are calculated with respect to the known pdf q, and we have used Bayes' theorem to expand the posterior pdf $p(\mathbf{m}|\mathbf{d}_{\text{obs}})$ in the second line of equation (3). It can be shown that the KL divergence is non-negative and only equals zero when $q(\mathbf{m}) = p(\mathbf{m}|\mathbf{d}_{\text{obs}})$ (Kullback and Leibler, 1951). Because the evidence term $\log p(\mathbf{d}_{\text{obs}})$ is computationally intractable, the KL divergence cannot be calculated directly. We therefore rearrange the above equation by moving the evidence term and the KL divergence onto the same side:

$$\begin{aligned}&\log p(\mathbf{d}_{\text{obs}}) - \mathrm{KL}[q(\mathbf{m})||p(\mathbf{m}|\mathbf{d}_{\text{obs}})] \\ &\quad= \mathrm{E}_q[\log p(\mathbf{m},\mathbf{d}_{\text{obs}})] - \mathrm{E}_q[\log q(\mathbf{m})]\end{aligned} \tag{4}$$

Given that the KL divergence is non-negative, the left-hand side defines a lower bound on the evidence, which is therefore called the evidence lower bound (ELBO):

$$\begin{aligned}\mathrm{ELBO}[q] &= \log p(\mathbf{d}_{\text{obs}}) - \mathrm{KL}[q(\mathbf{m})||p(\mathbf{m}|\mathbf{d}_{\text{obs}})] \\ &= \mathrm{E}_q[\log p(\mathbf{m},\mathbf{d}_{\text{obs}})] - \mathrm{E}_q[\log q(\mathbf{m})]\end{aligned} \tag{5}$$

The latter epxression can be estimated in practice using numerical methods because it does not involve the intractable evidence term. Since the evidence $\log p(\mathbf{d}_{\text{obs}})$ is a constant for a specific problem, minimizing the KL-divergence is equivalent to maximizing the ELBO. Variational inference in equation (2) can therefore be expressed as:

$$q^*(\mathbf{m}) = \arg\max_{q\in Q} \mathrm{ELBO}[q(\mathbf{m})] \tag{6}$$

In variational inference, the choice of the variational family Q is important because it determines both the accuracy of the approximation and the complexity of the optimisation problem. A good choice should be a family which is rich enough to approximate the posterior pdf accurately or at least provides the information that we seek, but simple enough such that the optimisation problem is tractable. Different choices of family may also allow different types of algorithm to be developed. In the VIP package we implement three different algorithms, ADVI, SVGD and sSVGD to solve inverse problems.

2.2 Automatic differential variational inference (ADVI)

ADVI uses the family of (transformed) Gaussians to solve variational inference problems (Kucukelbir et al., 2017). The transform arises because physical model parameters describe quantities that often have hard bounds, while Gaussian variables have infinite support. We therefore first transform the physical parameters into an unconstrained space using an invertible transform $T : \boldsymbol{\theta} = T(\mathbf{m})$. In this unconstrained space the joint distribution $p(\mathbf{m}, \mathbf{d}_{\text{obs}})$ becomes:

$$p(\boldsymbol{\theta}, \mathbf{d}_{\text{obs}}) = p(\mathbf{m}, \mathbf{d}_{\text{obs}})|\det \mathbf{J}_{T^{-1}}(\boldsymbol{\theta})| \tag{7}$$

where $\mathbf{J}_{T^{-1}}(\boldsymbol{\theta})$ is the Jacobian matrix of the inverse of T which accounts for the effects of changes in hypervolume between the unconstrained and constrained parameter spaces. In this unconstrained space define a Gaussian variational family

$$q(\boldsymbol{\theta}; \boldsymbol{\zeta}) = \mathcal{N}(\boldsymbol{\theta}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{8}$$

where $\boldsymbol{\zeta}$ represents variational parameters, that is, the mean vector $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$. To ensure that the covariance matrix $\boldsymbol{\Sigma}$ is positive semi-definite, we use a Cholesky factorization $\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}^{\mathrm{T}}$ where $\mathbf{L}$ is a lower triangular matrix, to reparameterise $\boldsymbol{\Sigma}$.

With the above definition, the variational problem in equation (6) becomes:

$$\begin{aligned}\boldsymbol{\zeta}^* &= \arg\max_{\boldsymbol{\zeta}} \text{ELBO}[q(\boldsymbol{\theta}; \boldsymbol{\zeta})] \\ &= \arg\max_{\boldsymbol{\zeta}} \mathrm{E}_q[\log p(\boldsymbol{\theta}, \mathbf{d}_{\text{obs}})] - \mathrm{E}_q[\log q(\boldsymbol{\theta}; \boldsymbol{\zeta})] \\ &= \arg\max_{\boldsymbol{\zeta}} \mathrm{E}_q[\log p\big(T^{-1}(\boldsymbol{\theta}), \mathbf{d}_{\text{obs}}\big) + \log|det \mathbf{J}_{T^{-1}}(\boldsymbol{\theta})|] \\ &\quad - \mathrm{E}_q[\log q(\boldsymbol{\theta}; \boldsymbol{\zeta})]\end{aligned} \tag{9}$$

This optimisation problem can be solved by using a gradient ascent algorithm. As shown in Kucukelbir et al. (2017), the gradients of the ELBO with respect to variational parameters $\boldsymbol{\mu}$ and $\mathbf{L}$ can be calculated using:

$$\begin{aligned}\nabla_{\boldsymbol{\mu}}\text{ELBO} = \mathrm{E}_{N(\boldsymbol{\eta}|\mathbf{0},\mathbf{I})}\big[\nabla_{\mathbf{m}} \log p(\mathbf{m}, \mathbf{d}_{\text{obs}}) \nabla_{\boldsymbol{\theta}} T^{-1}(\boldsymbol{\theta}) \\ + \nabla_{\boldsymbol{\theta}} \log|det \mathbf{J}_{T^{-1}}(\boldsymbol{\theta})|\big]\end{aligned} \tag{10}$$

$$\begin{aligned}\nabla_{\mathbf{L}}\text{ELBO} = \mathrm{E}_{N(\boldsymbol{\eta}|\mathbf{0},\mathbf{I})}\big[\big(\nabla_{\mathbf{m}} \log p(\mathbf{m}, \mathbf{d}_{\text{obs}}) \nabla_{\boldsymbol{\theta}} T^{-1}(\boldsymbol{\theta}) \\ + \nabla_{\boldsymbol{\theta}} \log|det \mathbf{J}_{T^{-1}}(\boldsymbol{\theta})|\big)\boldsymbol{\eta}^{\mathrm{T}}\big] + (\mathbf{L}^{-1})^{\mathrm{T}}\end{aligned} \tag{11}$$

where $\boldsymbol{\eta}$ is a random variable distributed according to the standard normal distribution $N(\boldsymbol{\eta}|\mathbf{0}, \mathbf{I})$. The expectations can be estimated using Monte Carlo (MC) integration, which in practice only requires a low number of samples because the optimisation is performed over many iterations so that statistically the gradients will lead to convergence towards the correct solution (Kucukelbir et al., 2017). The variational problem in equation (9) can now be solved by using gradient ascent methods. In the VIP package we implement four optimisation algorithms: stochastic gradient descent (SGD), ADAGRAD (Duchi et al., 2011), ADADELTA (Zeiler, 2012) and ADAM (Kingma and Ba, 2014). The final approximation to the Bayesian solution can be obtained by transforming $q(\boldsymbol{\theta}; \boldsymbol{\zeta}^*)$ back to the original space.

For transform T we implement a commonly-used logarithmic transform (Team et al., 2016; Zhang and Curtis, 2020a)

$$\begin{aligned}\theta_i &= T(m_i) = \log(m_i - a_i) - \log(b_i - m_i) \\ m_i &= T^{-1}(\theta_i) = a_i + \frac{(b_i - a_i)}{1 + exp(-\theta_i)}\end{aligned} \tag{12}$$

where m_i and θ_i represent the i^{th} parameter in the original and transformed space respectively, and a_i and b_i are the lower and upper bound on m_i. The final approximation obtained using ADVI is therefore limited in complexity by the Gaussian distribution $q(\boldsymbol{\theta}; \boldsymbol{\zeta}^*)$ and the transform T. Note that if no transform is performed, the method approximates the posterior pdf using a Gaussian distribution directly.

2.3 Stein variational gradient descent (SVGD)

Instead of using a specific form of pdf (for example, the Gaussian distribution in ADVI) in variational inference, it is also possible to use the density of a set of samples to represent the approximating probability distribution. SVGD is one such method in which the set of samples are called particles. In SVGD those particles are iteratively updated by minimizing the KL divergence so that the density of the final set of particles is distributed according to the posterior probability distribution. If we define the set of particles as $\{\mathbf{m}_i\}$, SVGD updates each particle using a smooth transform:

$$T(\mathbf{m}_i) = \mathbf{m}_i + \epsilon\boldsymbol{\phi}(\mathbf{m}_i) \tag{13}$$

where $\mathbf{m}_i$ is the i^{th} particle, $\boldsymbol{\phi}(\mathbf{m}_i)$ is a smooth vector function which describes the perturbation direction, and ϵ is the magnitude of the perturbation. When ϵ is sufficiently small, the transform is invertible since the Jacobian of the transform is close to an identity matrix. Denote $q(\mathbf{m})$ as the pdf represented by the set of particles, and $q_T(\mathbf{m})$ as the transformed probability distribution of $q(\mathbf{m})$ using equation (13). In order to reduce the KL divergence between $q_T(\mathbf{m})$ and $p(\mathbf{m}|\mathbf{d}_{\text{obs}})$, we first calculate the gradient of the KL divergence with respect to ϵ, which is found to be (Liu and Wang, 2016):

$$\nabla_{\epsilon}\text{KL}[q_T||p]\,|_{\epsilon=0} = -\mathrm{E}_q\left[trace\left(\mathcal{A}_p\boldsymbol{\phi}(\mathbf{m})\right)\right] \tag{14}$$

where $\mathcal{A}_p$ is the Stein operator defined as $\mathcal{A}_p\boldsymbol{\phi}(\mathbf{m}) = \nabla_{\mathbf{m}}\log p(\mathbf{m}|\mathbf{d}_{\text{obs}})\boldsymbol{\phi}(\mathbf{m})^T + \nabla_{\mathbf{m}}\boldsymbol{\phi}(\mathbf{m})$. This equation implies that one can obtain the steepest descent direction of the KL-divergence by maximizing the right-hand expectation $\mathrm{E}_q\left[trace\left(\mathcal{A}_p\boldsymbol{\phi}(\mathbf{m})\right)\right]$, and consequently the KL divergence can be reduced by stepping a small distance in that direction. Iteratively re-calculating equation (14) and stepping in each revised direction locates a minimum in the KL divergence.

The optimal direction $\boldsymbol{\phi}^*$ that maximizes the expectation $\mathrm{E}_q\left[trace\left(\mathcal{A}_p\boldsymbol{\phi}(\mathbf{m})\right)\right]$ in equation (14) can be found using kernels. Assume $x, y \in X$ and define a mapping ϕ from X to a space where an inner product $\langle, \rangle$ is defined (called a Hilbert space); a *kernel* is a function that satisfies $k(x, y) = \langle\phi(x), \phi(y)\rangle$. Given a kernel function $k(\mathbf{m}', \mathbf{m})$, the optimal $\boldsymbol{\phi}^*$ can be calculated using (see details in Liu and Wang, 2016):

$$\boldsymbol{\phi}^* \propto \mathrm{E}_{\{\mathbf{m}' \sim q\}}[\mathcal{A}_p k(\mathbf{m}', \mathbf{m})] \tag{15}$$

In the VIP package, we implement a commonly-used kernel function, the radial basis function (RBF):

$$k(\mathbf{m},\mathbf{m}') = \exp[-\frac{\|\mathbf{m}-\mathbf{m}'\|^2}{2h^2}] \tag{16}$$

where h is a scale factor that controls the magnitude of similarity between the two particles based on their distance apart. Given equations (14) and (15), the KL divergence can be minimized by iteratively applying the transform in equation (13) with the optimal $\boldsymbol{\phi}^*$ to a set of initial particles:

$$\mathbf{m}_i^{l+1} = T(\mathbf{m}_i^l) = \mathbf{m}_i^l + \epsilon^l \boldsymbol{\phi}_l^*(\mathbf{m}_i^l) \tag{17}$$

where l represents the l^{th} iteration. Note that the expectation in equation (15) can be estimated using the particles' mean value, so we can compute $\boldsymbol{\phi}_l^*$ using:

$$\begin{aligned}\boldsymbol{\phi}_l^*(\mathbf{m}) &= \frac{1}{n}\sum_{j=1}^{n}\left[\mathcal{A}_p k(\mathbf{m}_j^l,\mathbf{m})\right] \\ &= \frac{1}{n}\sum_{j=1}^{n}\left[k(\mathbf{m}_j^l,\mathbf{m})\nabla_{\mathbf{m}_j^l}\text{log}p(\mathbf{m}_j^l|\mathbf{d}_{\text{obs}})\right. \\ &\left.+\nabla_{\mathbf{m}_j^l}k(\mathbf{m}_j^l,\mathbf{m})\right]\end{aligned} \tag{18}$$

where n is the number of particles. For sufficiently small $\{\epsilon^l\}$ the transform is invertible, and the process converges to the posterior distribution asymptotically as $n \to \infty$ (Liu and Wang, 2016). Note that even though the posterior distribution $p(\mathbf{m}_j^l|\mathbf{d}_{\text{obs}})$ is unknown in practice, we can always calculate its value up to an unknown constant for a specific model. As a result, its gradient $\nabla_{\mathbf{m}_j^l}\text{log}p(\mathbf{m}_j^l|\mathbf{d}_{\text{obs}})$ can be obtained, and hence the $\boldsymbol{\phi}_l^*$.

The first term in equation (15) is the kernel weighted average of gradients of the posterior pdf from all particles, and drives particles toward high probability areas. For the RBF kernel the second term becomes $\sum_j \frac{\mathbf{m}-\mathbf{m}_j}{\sigma^2}k(\mathbf{m}_j,\mathbf{m})$ which move particles away from its neighbouring particles. This term therefore acts as a repulsive force that prevents particles from collapsing to a single mode. SVGD balances the drive towards high probabilities and the repulsive force such that the density of particles moves towards the posterior pdf.

Note that the scale factor h in the RBF kernel controls the weighting value of particles. As suggested in several studies (Liu and Wang, 2016; Zhang and Curtis, 2020a), we take h as $\tilde{d}/\sqrt{2\text{log}n}$ where $\tilde{d}$ is the median of pairwise distances between all particles. This choice enables that for particle $\mathbf{m}_i$ the contribution form its own gradient is balanced from all other particles as $\sum_{j\neq i}k(\mathbf{m}_i,\mathbf{m}_j) \approx n\text{exp}(-\frac{1}{2h^2}\tilde{d}^2) = 1$. If $h \to 0$, the method reduces to independent gradient ascent for each particle.

In SVGD the accuracy of estimation increases with the number of particles. For one single particle the method becomes a standard gradient ascent method toward the model with maximum a posterior (MAP) pdf value. This implies that even for a small number of particles SVGD can still produce an accurate parameter estimate as MAP estimation has been demonstrated to be an effective method in practice. Thus, in practice, one can start from a small number of particles and gradually increase the particles to produce more accurate estimates of the uncertainty.

2.4 Stochastic SVGD

Although SVGD has been applied in many fields (Gong et al., 2019; Zhang and Curtis, 2020a; Pinder et al., 2020; Ramgraber et al., 2021; Ahmed et al., 2022), the method can produce biased results in high dimensional problems because of the finite number of particles and the limitation of computational cost in practice (Ba et al., 2022). In order to further improve accuracy of the method, Gallego and Insua (2018) proposed a variant of SVGD, called stochastic SVGD (sSVGD), which combines SVGD and McMC by adding a Gaussian noise term to the dynamics of SVGD. By doing this sSVGD becomes an McMC method with multiple interacting Markov chains, and since every set of particle values can be regarded as a sample of the posterior pdf, the method can generate many samples that are distributed according to the posterior pdf. Under certain conditions (see below), sSVGD guarantees asymptotic convergence to the posterior pdf as the number of iterations tends to infinity, which standard SVGD with a finite number of particles cannot achieve. As a result sSVGD can produce more accurate results than the SVGD method, provided that the number of iterations is sufficient to remove effects of the distribution of samples near the start of the chain (the so-called burn-in period) (Gallego and Insua, 2018; Zhang et al., 2023).

To introduce sSVGD, we start from a stochastic differential equation (SDE). For a random variable $\mathbf{z}$, the SDE is defined as:

$$d\mathbf{z} = \mathbf{f}(\mathbf{z})dt + \sqrt{2\mathbf{D}(\mathbf{z})}d\mathbf{W}(t) \tag{19}$$

where $\mathbf{f}(\mathbf{z})$ is called the drift, $\mathbf{W}(t)$ is a Wiener process, and $\mathbf{D}(\mathbf{z})$ represents a positive semidefinite diffusion matrix. All continuous Markov process can be expressed as an SDE, and consequently one can construct a Markov chain by simulating the SDE (Oksendal, 2013). Assume $p(\mathbf{z})$ as the posterior distribution, an SDE that converges to the $p(\mathbf{z})$ can be constructed as (Ma et al., 2015):

$$\mathbf{f}(\mathbf{z}) = [\mathbf{D}(\mathbf{z}) + \mathbf{Q}(\mathbf{z})]\,\nabla\text{log}p(\mathbf{z}) + \Gamma(\mathbf{z}) \tag{20}$$

where $\mathbf{Q}(\mathbf{z})$ is a skew-symmetric curl matrix, and $\Gamma_i(\mathbf{z}) = \sum_{j=1}^{d}\frac{\partial}{\partial \mathbf{z}_j}(\mathbf{D}_{ij}(\mathbf{z}) + \mathbf{Q}_{ij}(\mathbf{z}))$. To simulate this process, we can discretize the above equation using the Euler-Maruyama discretization:

$$\begin{aligned}\mathbf{z}_{t+1} = \mathbf{z}_t &+ \epsilon_t\left[(\mathbf{D}\left(\mathbf{z}_t\right) + \mathbf{Q}(\mathbf{z}_t))\,\nabla\text{log}p(\mathbf{z}_t) + \Gamma(\mathbf{z}_t)\right] \\ &+ \mathcal{N}(\mathbf{0}, 2\epsilon_t\mathbf{D}(\mathbf{z}_t))\end{aligned} \tag{21}$$

where $\mathcal{N}(\mathbf{0}, 2\epsilon_t\mathbf{D}(\mathbf{z}_t))$ represents a Gaussian distribution with covariance $2\epsilon_t\mathbf{D}(\mathbf{z}_t)$. The gradient $\nabla\text{log}p(\mathbf{z}_t)$ can be computed using the full data set, or using uniformly randomly selected minibatch data subsets which results in a stochastic gradient approximation. In either case the above process converges to the posterior distribution asymptotically as $\epsilon_t \to 0$ and $t \to \infty$ (Ma et al., 2015). Matrices $\mathbf{D}(\mathbf{z})$ and $\mathbf{Q}(\mathbf{z})$ can be adjusted to obtain faster convergence to the posterior distribution. For example, if we set $\mathbf{D}=\mathbf{I}$ and $\mathbf{Q}=\mathbf{0}$, one obtains stochastic gradient Langevin dynamics (Welling and Teh, 2011). If we construct an augmented space $\overline{\mathbf{z}}=(\mathbf{z},\mathbf{x})$ by concatenating a moment term $\mathbf{x}$ to the state space $\mathbf{z}$, and set $\mathbf{D}=\mathbf{0}$ and $\mathbf{Q}=\begin{pmatrix}\mathbf{0} & -\mathbf{I}\\ \mathbf{I} & \mathbf{0}\end{pmatrix}$ then the stochastic Hamiltonian Monte Carlo method can be derived (Chen et al., 2014).

In sSVGD we define an augmented space $\mathbf{z}=(\mathbf{m}_1,\mathbf{m}_2,...,\mathbf{m}_n)$ by concatenating the set of particles $\{\mathbf{m}_i\}$, and use equation (21) to generate samples from the posterior distribution $p(\mathbf{z})=\prod_{i=1}^{n}p(\mathbf{m}_i|\mathbf{d}_{\text{obs}})$. Define a matrix $\mathbf{K}$

$$\mathbf{K}=\frac{1}{n}\begin{bmatrix}k(\mathbf{m}_1,\mathbf{m}_1)\mathbf{I}_{d\times d} & \dots & k(\mathbf{m}_1,\mathbf{m}_n)\mathbf{I}_{d\times d}\\ \vdots & \ddots & \vdots\\ k(\mathbf{m}_n,\mathbf{m}_1)\mathbf{I}_{d\times d} & \dots & k(\mathbf{m}_n,\mathbf{m}_n)\mathbf{I}_{d\times d}\end{bmatrix} \tag{22}$$

where $k(\mathbf{m}_i,\mathbf{m}_j)$ is a kernel function defined in equation (16) and $\mathbf{I}_{d\times d}$ is an identity matrix. According to the definition of kernel functions, the matrix $\mathbf{K}$ is positive definite (Gallego and Insua, 2018). By setting $\mathbf{Q}(\mathbf{z}_t)=\mathbf{0}$ and $\mathbf{D}(\mathbf{z}_t)=\mathbf{K}$, we obtain the stochastic SVGD algorithm:

$$\mathbf{z}_{t+1}=\mathbf{z}_t+\epsilon_t[\mathbf{K}\nabla\text{log}p(\mathbf{z}_t)+\nabla\cdot\mathbf{K}]+\mathcal{N}(\mathbf{0},2\epsilon_t\mathbf{K}) \tag{23}$$

Note that without the noise term $\mathcal{N}(\mathbf{0},2\epsilon_t\mathbf{K})$, the above equation becomes the standard SVGD method – compare equations (23) with equation (18), repeated here:

$$\mathbf{z}_{t+1}=\mathbf{z}_t+\epsilon_t[\mathbf{K}\nabla\text{log}p(\mathbf{z}_t)+\nabla\cdot\mathbf{K}] \tag{24}$$

sSVGD is therefore an McMC method that uses the gradients from SVGD to produce successive samples. According to equation (20), this process converges to $p(\mathbf{z})=\prod_{i=1}^{n}p(\mathbf{m}_i|\mathbf{d}_{\text{obs}})$ asymptotically. Note that when n is sufficiently large, the noise term $\mathcal{N}(\mathbf{0},2\epsilon_t\mathbf{K})$ becomes arbitrarily small. In such cases sSVGD and SVGD produce the same results.

The process defined in equation (23) requires samples to be generated from the distribution $\mathcal{N}(\mathbf{0},2\epsilon_t\mathbf{K})$. In order to perform this efficiently, we first define a matrix $\mathbf{D}_{\mathbf{K}}$

$$\mathbf{D}_{\mathbf{K}}=\frac{1}{n}\begin{bmatrix}\overline{\mathbf{K}} & & \\ & \ddots & \\ & & \overline{\mathbf{K}}\end{bmatrix} \tag{25}$$

where $\overline{\mathbf{K}}$ is an $n\times n$ matrix with $\overline{\mathbf{K}}_{ij}=k(\mathbf{m}_i,\mathbf{m}_j)$. The matrix $\mathbf{D}_{\mathbf{K}}$ can be constructed from $\mathbf{K}$ using $\mathbf{D}_{\mathbf{K}}=\mathbf{P}\mathbf{K}\mathbf{P}^{\text{T}}$ where $\mathbf{P}$ is a permutation matrix

$$\mathbf{P}=\left[\begin{array}{c|c|c|c}
1 & & & \\
 & 1 & & \\
 & & \ddots & \\
 & & & 1\\ \hline
1 & & & \\
 & 1 & & \\
 & & \ddots & \\
 & & & 1\\ \hline
\ddots & \ddots & \ddots & \ddots\\ \hline
1 & & & \\
 & 1 & & \\
 & & \ddots & \\
 & & & 1
\end{array}\right] \tag{26}$$

The action of this permutation matrix on a vector $\mathbf{z}$ rearranges the order of the vector from the basis where the particles are listed sequentially to that where the first coordinates of all particles are listed, then the second, etc. With these definitions, a random sample $\boldsymbol{\eta}$ can be generated efficiently using

$$\begin{aligned}\boldsymbol{\eta}&\sim\mathcal{N}(\mathbf{0},2\epsilon_t\mathbf{K})\\ &\sim\sqrt{2\epsilon_t}\mathbf{P}^{\text{T}}\mathbf{P}\mathcal{N}(\mathbf{0},\mathbf{K})\\ &\sim\sqrt{2\epsilon_t}\mathbf{P}^{\text{T}}\mathcal{N}(\mathbf{0},\mathbf{D}_{\mathbf{K}})\\ &\sim\sqrt{2\epsilon_t}\mathbf{P}^{\text{T}}\mathbf{L}_{\mathbf{D}_{\mathbf{K}}}\mathcal{N}(\mathbf{0},\mathbf{I})\end{aligned} \tag{27}$$

where $\mathbf{L}_{\mathbf{D}_{\mathbf{K}}}$ is the lower triangular Cholesky decomposition of matrix $\mathbf{D}_{\mathbf{K}}$. Taking into account the fact that $\mathbf{D}_{\mathbf{K}}$ is a block-diagonal matrix, $\mathbf{L}_{\mathbf{D}_{\mathbf{K}}}$ can be computed easily as only the lower triangular Cholesky decomposition of matrix $\overline{\mathbf{K}}$ is required. In practice this calculation is computationally negligible because the number of particles n is usually modest (<1000). One can now use equation (23) to generate samples from the posterior distribution.

3 Code overview

The VIP package implements the suite of variational methods to solve geophysical inverse problems using the Python programming language. The package includes a set of specific forward and inverse problems such as 2D travel time tomography and 2D full waveform inversion, and also allows users to provide their own forward functions. In variational inference one needs to compute the gradient of the posterior pdf with respect to model parameters. We use the adjoint method to calculate the gradient in the case of seismic full waveform inversion (Lions, 1971; Tarantola, 1984; Tromp et al., 2005; Fichtner et al., 2006; Plessix, 2006), and the ray tracing method in the case of travel time tomography (Rawlinson and Sambridge, 2004). For user-specified forward problems it is required that users implement their own function that computes gradients.

The prior pdf is important in Bayesian inference as it provides information about model parameters independent of the data. The VIP package provides two commonly-used prior distributions: Uniform and Gaussian pdf's (note that these are only used as prior pdf's, and do not place any additional constraints on the variational families described above). To implement the Uniform distribution we employ two strategies. In the first strategy we impose hard constraints on model parameters, that is, for any parameter that assumes a value outside the distribution we reset the value to be the closest limit. Note that a similar strategy cannot be used in ADVI as the method assumes a Gaussian variational family which cannot be defined in a constrained space. The second strategy involves using equation (12) to transform model parameters into an unconstrained space and perform variational inversion in that space, which provides a more flexible way to employ a variety of variational families. In addition, users can provide their own prior distributions by implementing an appropriate pdf function (see details in the code documentation).

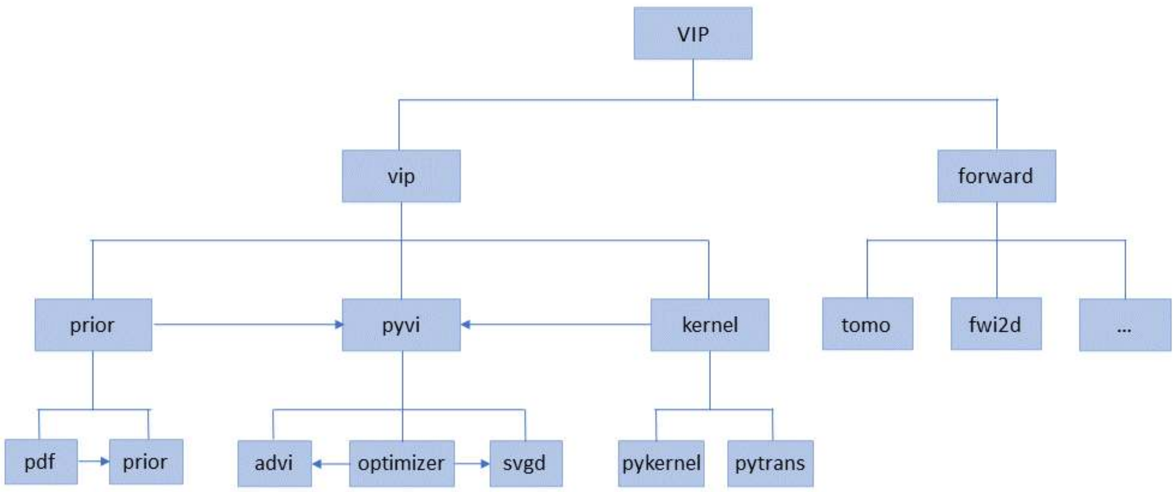


Figure 1 Code structure of VIP. Each rectangle represents a folder or file in the package. Users can implement their own forward functions similarly to the way this is implemented in examples *tomo* and *fwi2d*.

Python is a popular high-level interpreted programming language which suffers from slow execution for computationally intensive numerical simulations. We therefore implement time-consuming components of the code (e.g., the forward modelling functions) using Fortran and produce compiled C extensions for these codes using the Cython framework (Behnel et al., 2010). By doing this the code achieves C-like speeds. To further improve efficiency of the code, we use a Python library called Dask, which is designed for parallel and distributed computing, to parallelize the forward computation at the sample (particle) level (Rocklin et al., 2015). The package therefore provides an efficient, scalable and user-friendly implementation which can be deployed on a desktop as well as modern high performance computation facilities. Our aim is to implement a framework which can be used to solve various inverse problems, ranging from educational examples to complex, realistic studies.

Figure 1 shows the structure of VIP. The inversion code (*vip* in Figure 1) is implemented separately from forward modelling codes (*forward* in Figure 1), and only requires an interface of forward functions that returns logarithmic posterior pdf values and gradients (details can be found in the code documentation and in two examples *tomo* and *fwi2d*). Thus, users can easily combine their own forward functions with the package. In the *vip* code the prior distributions, kernel functions and variational algorithms are implemented in three different directories (*prior*, *kernel* and *pyvi* in Figure 1) so that the code can easily be extended to other prior pdfs, kernel functions and variational methods. For example, users can implement their own prior pdfs by adding a proper pdf function in the *pdf* code in the *prior* directory. Note that both SVGD and sSVGD methods are implemented in the *svgd* code.

4 Applications

4.1 Travel time tomography

As a first example we use the VIP package to solve a 2D tomographic problem. Specifically, we create Love wave group velocity maps of the British Isles using ambient seismic noise data recorded by 61 seismometers (blue triangles in Figure 2a). The geological setting and the main terrain boundaries of the British Isles are shown in Figure 2b. The ambient noise data were recorded in 2001-2003, 2006-2007 and in 2010 using three different subarrays. The two horizontal components of the data (N and E) were first rotated to the transverse and radial directions, and the obtained transverse data were cross correlated to produce Love waves between different station pairs. Travel times associated with group velocity at different periods between different station pairs are then estimated from those love waves. Details of the data processing procedures can be found in (Galetti et al., 2017). In this study we use a total number of 401 travel time measurements at 10 s period.

We parameterise the study region using a regular grid of 37×40 cells with a spacing of $0.33°$in both longitude and latitude directions. The prior pdf for group velocity in each cell is set to be a Uniform distribution between 1.56 km/s to 4.8 km/s, of which the lower and upper bound were chosen to exceed the range of group velocities between all station pairs when assuming a great circle ray path (Zhao et al., 2022a). The likelihood function is chosen to be a Gaussian distribution to represent the data noise, which is estimated from independent travel time measurements by stacking randomly selected subsets of daily cross correlations (Galetti et al., 2017). In the inversion the predicted travel times are calculated using the fast marching method (Rawlinson and Sambridge, 2004).

We apply the above suite of methods to solve this tomographic problem, and compare the results with those obtained using the Metropolis-Hastings McMC

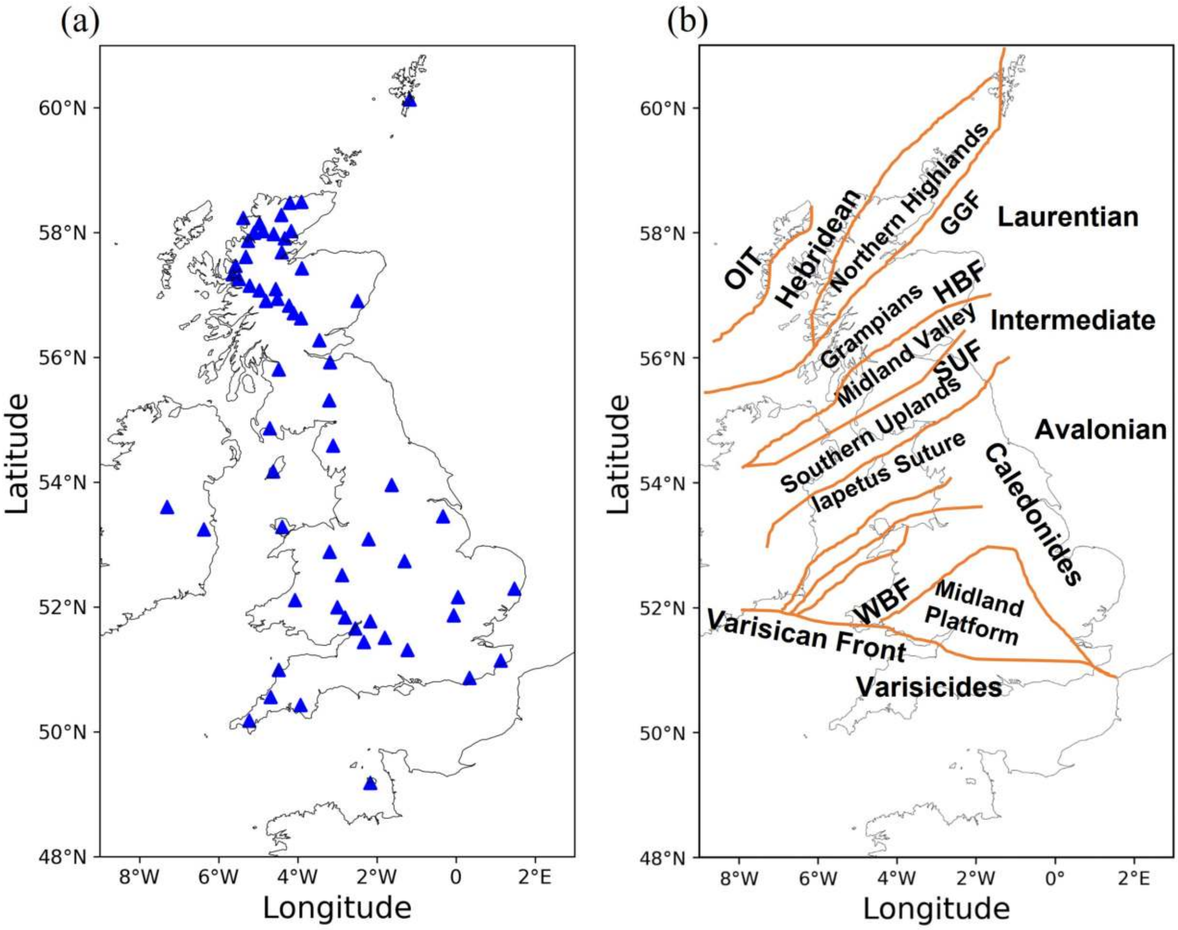


Figure 2 **(a)** Locations of seismometers (blue triangles) around British Isles used in this study. **(b)** Terrane boundaries in the British Isles from Galetti et al. (2017). Abbreviations are as follows: OIT, Outer Isles Thrust; GGF, Great Glen Fault; HBF, Highland Boundary Fault; SUF, Southern Uplands Fault; WBF, Welsh Borderland Fault System.

(MH-McMC) method (Zhao et al., 2022a). The Uniform prior distribution is implemented using the second strategy that transforms variables into an unconstrained space in variational inversions. For ADVI, we started the method with a standard Gaussian distribution in the unconstrained space, and performed 10,000 iterations at which point the misfit value ceases to decrease using the ADAM optimisation algorithm (Kingma and Ba, 2014). To visualize the results we generated 5,000 samples from the obtained Gaussian distribution and transformed them back to the original space to estimate posterior statistics. For SVGD, we generated 500 particles from the prior distribution and updated them using equation (18) for 3,000 iterations at which point the mean and standard deviation models became stable. The final particles are used to calculate the mean and standard deviation of the posterior distribution. For sSVGD, we started from 20 particles generated from the prior distribution, and updated them using equation (23) for 6,000 iterations after an additional burn-in period of 2,000 iteration, after which the average misfit value across all particles became approximately stationary. To reduce the memory and storage cost, we only retained samples every fourth iteration after the burn-in period, which results in a total of 30,000 samples.

Figure 3 shows the mean and standard deviation maps obtained using the suite of variational methods, as well as those obtained using the MH-McMC algorithm (Zhao et al., 2022a). Overall the results obtained using different methods show similar mean structures which have a good agreement with the known geology and previous tomographic studies in the British Isles (Nicolson et al., 2012, 2014; Galetti et al., 2017; Zhao et al., 2022a). For example, in the Scottish highlands the mean maps clearly exhibit high velocities (annotation 1 in Figure 3) which are consistent with the distribution of Lewisian and Dalradian complexes in this area. Similarly high velocities associated with the accretionary complex of the Southern Uplands (annotation 2) are clearly visible around 4°W, 55°N following a SW-NE trend. Between the Highland Boundary Fault and the Southern Uplands Fault a similar trend of low velocity zone (annotation 3) is found in the Midland Valley. Low velocities are also observed in a number of sedimentary basins such as the East Irish Sea (4.5°W, 54°E - annotation 4), the Cheshire Basin (2.5°W, 52.5°E - annotation 6), the Anglian-London Basin (0°, 52°N - annotation 7), the Weald Basin (0°, 51°N - annotation 8) and the Wessex Basin (3°W, 50.5°N - annotation 9). By contrast, high velocities can be found in granitic intrusion regions, for example, in northwest Wales (around 4°W, 53°N - annotation 5) and Cornwall (around 4.5°W, 50.5°N - annotation 10). More detailed

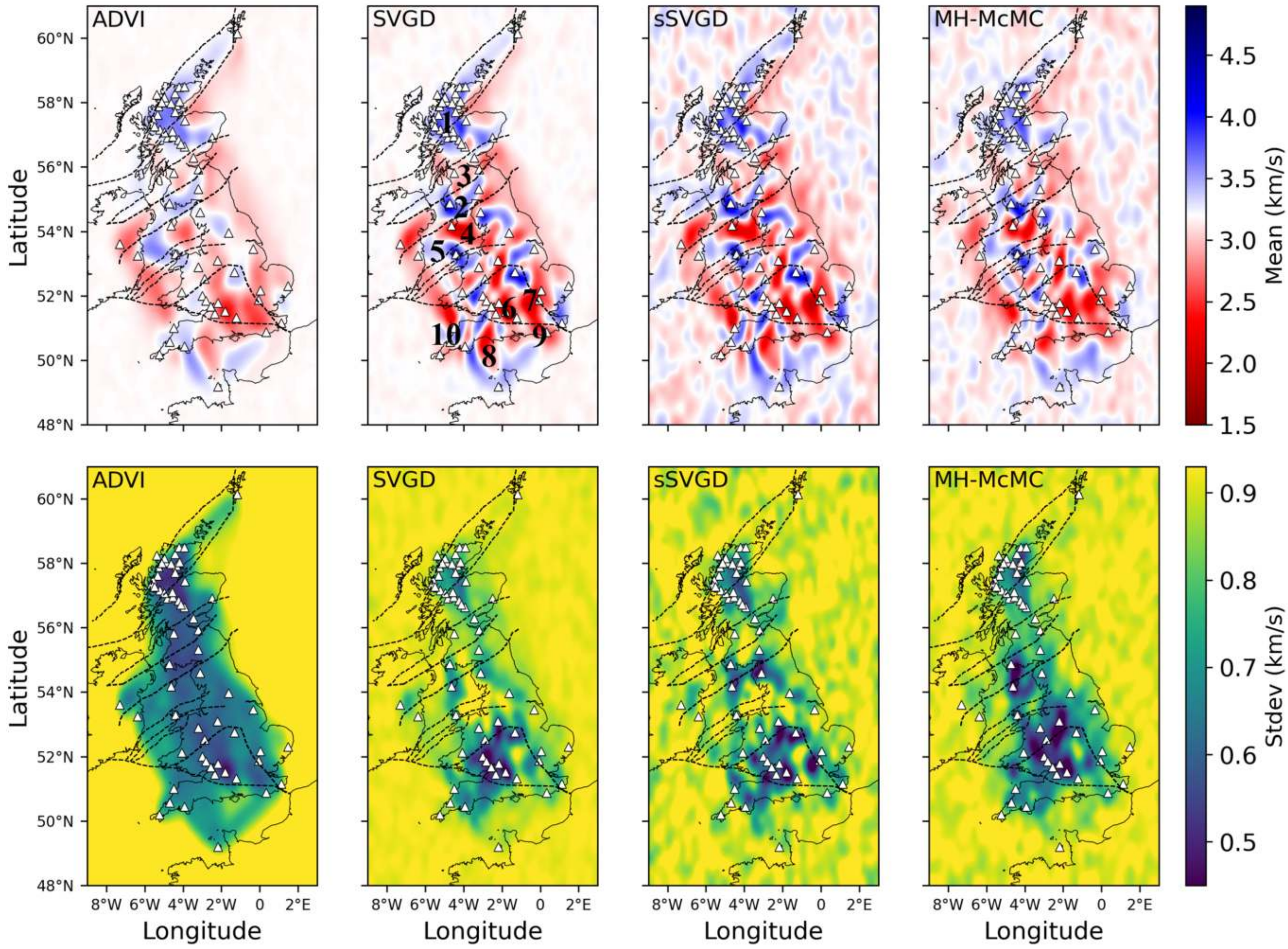


Figure 3 Mean (top row) and standard deviation (bottom row) maps of group velocity at 10 s period obtained using ADVI, SVGD, sSVGD and MH-McMC respectively. White triangles denote locations of seismometers. Black dashed lines show the Terrane boundaries in Figure 2. Black numbers are referred to in the main text.

discussion and interpretation of the velocity structures can be found in Galetti et al. (2017).

Among these results the mean map obtained using ADVI shows the smoothest structure, whereas other maps provide more detailed information. This has also been observed in previous studies (Zhang and Curtis, 2020a; Zhao et al., 2022a) and is likely caused by the limitation of implicit Gaussian assumption made in ADVI. In far offshore areas because few ray paths go through the open marine regions, the mean maps obtained using ADVI and SVGD show almost homogeneous velocity structure across these areas whose value is consistent with the mean of prior distribution. In comparison, the results obtained using sSVGD and MH-McMC exhibit more heterogeneous structures, which probably indicates that the two methods have not converged sufficiently. These areas are only loosely constrained by the data (or not at all) and hence have large posterior uncertainties requiring many more randomly generated samples in order to explore and represent the posterior distribution accurately compared to areas with tighter constraints from the data. Note that both sSVGD and MH-McMC involve random sampling of the posterior distribution, whereas samples obtained using SVGD are found deterministically by optimisation. As a result, SVGD produces smoother results (Zhang and Curtis, 2021; Zhang et al., 2023).

Overall the standard deviation maps obtained using SVGD, sSVGD and MH-McMC show similar structures. For example, the results show lower uncertainties in the Scottish highlands and southern England because of dense arrays in those areas. In the offshore areas the standard deviation is around 0.93 which is the standard deviation of the prior as no ray path goes through these regions. On the east side of the island just off the coast, although no seismometer is deployed, there are rays that travel through those areas (see details in Galetti et al., 2017), and consequently the standard deviation is smaller than that of the prior. There is a high uncertainty loop around the low velocity anomaly in the Anglian-London Basin (annotation 7 in Figure 3), which has also been observed in previous studies (Galetti et al., 2015, 2017) and reflects uncertainty in the shape of the anomaly. In addition, the East Irish Sea (annotation 4) shows high uncertainties. This is probably because few ray paths go through this area due to its lower velocity, and consequently the area is not well constrained by the data. By contrast, the standard deviation map obtained using ADVI shows different features. Although in the Scottish highlands the results still show lower uncertainty, the rest of the area within the receiver array has almost the same uncertainty level with little variation. In addition, in the West Irish Sea and the North Sea area between Northern Scotland and Shetland Islands the results show lower uncertainties which are not observed in the results obtained using other methods. This suggests that ADVI can produce biased results because of its underlying Gaussian assumption as found in previous studies (Zhang and Curtis, 2020a).

Table 1 compares the number of forward simulations required by each method to obtain these results, which provides a good metric of the computation cost as the forward simulation is the most computationally expensive component of each method. Note that the three variational methods require computation of derivatives of the posterior pdf with respect to model parameters, which adds computational cost compared with the MH-McMC method. In this travel time tomography example the derivatives are calculated using ray paths, which are traced through the computed travel time field. This calculation requires a computation equivalent to approximately 0.08 forward simulations. We therefore compute the equivalent number of simulations by multiplying the number of simulations required by the three variational methods by 1.08, which are shown in the third column in Table 1.

The results indicate that ADVI is apparently the most efficient method as it only requires 10,000 simulations, but we have demonstrated that the method probably produces biased results. SVGD demands the highest computational cost among the three variational methods, while sSVGD requires about 10 times fewer simulations than SVGD. This makes sSVGD a good choice for practical applications as noted in Zhang et al. (2023). Nevertheless, all three variational methods are significantly more efficient than the basic MH-McMC method implemented here as a bench-mark, which required 15 millions simulations in total with 10 independent parallel chains.

We note that the above comparison depends on subjective assessment of the point of convergence for each method, so the absolute number of simulations required by each method may not be entirely accurate (especially the number used for the MH-McMC algorithm). Nevertheless the comparison at least provides insights into the relative computational cost of each method. A more careful and thorough comparison between the same MH-McMC method and variational methods can be found in Zhao et al. (2022a) which again demonstrated that variational methods were computationally efficient.

4.2 Full-waveform inversion

For the second example we use the VIP package to solve a 2D full waveform inversion problem. The input model is selected to be a part of the Marmousi model (Figure 4a, Martin et al., 2006), and is discretized using a regular 120×200 grid with a spacing of 20 m. Ten sources are equally distributed at 20 m water depth (red stars in Figure 4), and 200 receivers are equally spaced at the depth of 360 m on the seabed across the horizontal extent of the model. We simulate the waveform data using a time-domain finite difference method with a Ricker wavelet of 10 Hz central frequency, and added Gaussian noise to the data whose standard deviation is set to be 2 percent of the median of the maximum amplitude of each seismic trace. The gradients of the logarithm posterior pdf with respect to velocity are calculated using the adjoint method (Tarantola, 1988; Tromp et al., 2005; Fichtner et al., 2006; Plessix, 2006).

The prior distribution is set to be a Uniform distribution over an interval of 2 km/s at each depth (Figure 4b). To ensure that the rock velocity is higher than the velocity in the water, we imposed an extra lower bound of 1.5 km/s. For the likelihood function we use a Gaussian distribution to represent uncertainties on the waveform data:

$$p(\mathbf{d}_{\text{obs}}|\mathbf{m}) \propto \exp\left[-\frac{1}{2}\sum_i \left(\frac{d_i^{\text{obs}} - d_i(\mathbf{m})}{\sigma_i}\right)^2\right] \tag{28}$$

where i is the index of time samples, and σ_i is the standard deviation of that sample.

We apply SVGD and sSVGD to solve this full waveform inversion problem as we have demonstrated that these methods provide more accurate results than ADVI. For SVGD we used 600 particles that are initially generated from the prior distribution (an example is shown in Figure 4c) and updated them using equation (18) for 600 iterations. The final particles are used to calculate statistics of the posterior distribution. For sSVGD we generated 20 particles from the prior distribution and updated them for 4,000 iterations after an additional burn-in period of 2,000. Similarly, to reduce the memory and storage cost we only retain samples from every tenth iterations, which results in a total of 8,000 samples. Those final samples are then used to compute statistics of the posterior distribution.

Figure 5 shows the mean and standard deviation models obtained using SVGD and sSVGD. Overall the two methods produce similar results. For example, both mean models (Figure 5a and c) show similar structures to the true structure, especially in the shallow part ($<$ 1.5 km). In the deep part ($>$ 1.5 km) and close to the sides, the mean models appear to be less similar to the true structure because the waveform data are less sensitive to the velocity structure in those areas. However, the mean obtained using sSVGD is more similar to the true structure than that obtained using SVGD. This reflects the fact that sSVGD can produce more accurate results than SVGD in high dimensional spaces, which has also been observed in other studies (Gallego and Insua, 2018; Zhang et al., 2023). Note that similarly to the travel time tomography example above, the mean obtained using SVGD shows smoother structures than that obtained using sSVGD. This is likely because sSVGD is an McMC method which generates samples using stochastic sam-

Method	Number of simulations	Comparable number of simulations
ADVI	10,000	10,800
SVGD	1500,000	1620,000
sSVGD	160,000	172,800
MH-McMC	15,000,000	15,000,000

Table 1 A comparison of computational cost for ADVI, SVGD, sSVGD and MH-McMC.

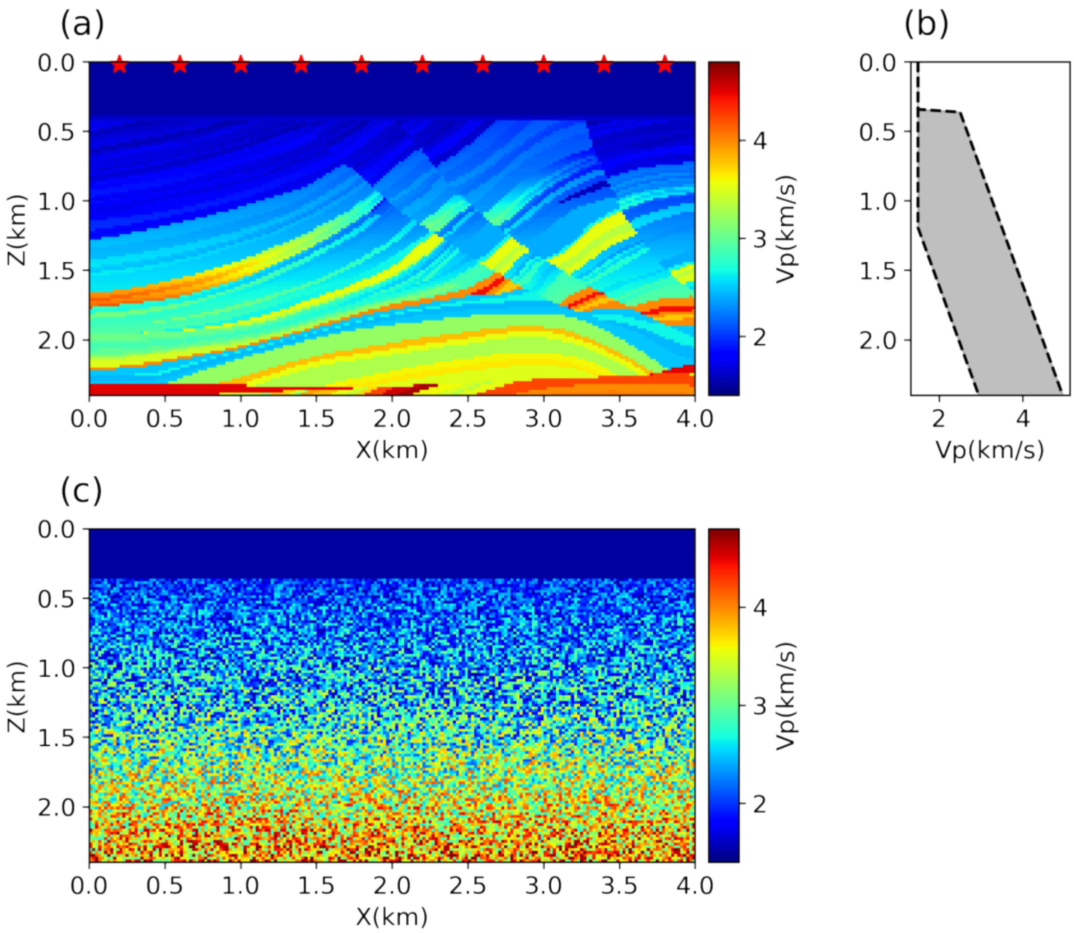


Figure 4 **(a)** The true structure used in the full waveform inversion example. Ten sources are located at the depth of 20 m (red stars) and 200 receivers (not shown) are equally spaced at the depth of 360 m on the seabed. **(b)** The prior distribution of seismic velocity, which is set to be a Uniform distribution with an interval of 2 km/s at each depth. An additional lower bound of 1.5 km/s is also imposed on the velocity to ensure that the rock velocity is higher than the velocity in water. **(c)** An example particle generated from the prior distribution.

pling, whereas in SVGD particles are obtained deterministically using optimisation. A similar phenomenon has also been observed in other studies when comparing results obtained using SVGD and sSVGD or McMC (Zhang and Curtis, 2021; Zhang et al., 2023).

Overall the standard deviation models show similar structural shapes to those in the mean model as has been observed in other studies (Gebraad et al., 2020; Zhang and Curtis, 2020b, 2021; Zhang et al., 2023). In the shallow part (< 1.0 km) the results show lower uncertainties and in the deeper part the uncertainty is higher because of lower data coverage. Those higher velocity anomalies in the deeper part are clearly associated with lower standard deviations, which likely reflects that those anomalies have large influences on the waveform data and hence have lower uncertainty. Similarly to the mean structures, the standard deviations obtained using SVGD show smoother structures than are obtained using sSVGD. In addition, the magnitude of the standard deviation obtained using SVGD is slightly lower than that obtained using sSVGD, which is likely because SVGD can underestimate uncertainties in high dimensional spaces due to the limited number of posterior samples produced (Ba et al., 2022; Zhang et al., 2023).

To further understand the results we show marginal distributions obtained using SVGD and sSVGD along three vertical profiles whose locations are denoted by dashed black lines in Figure 5. Overall the results show broader distributions in the deeper part (> 1 km) than in the shallow part as we have observed in the standard deviation models. Furthermore, the distributions

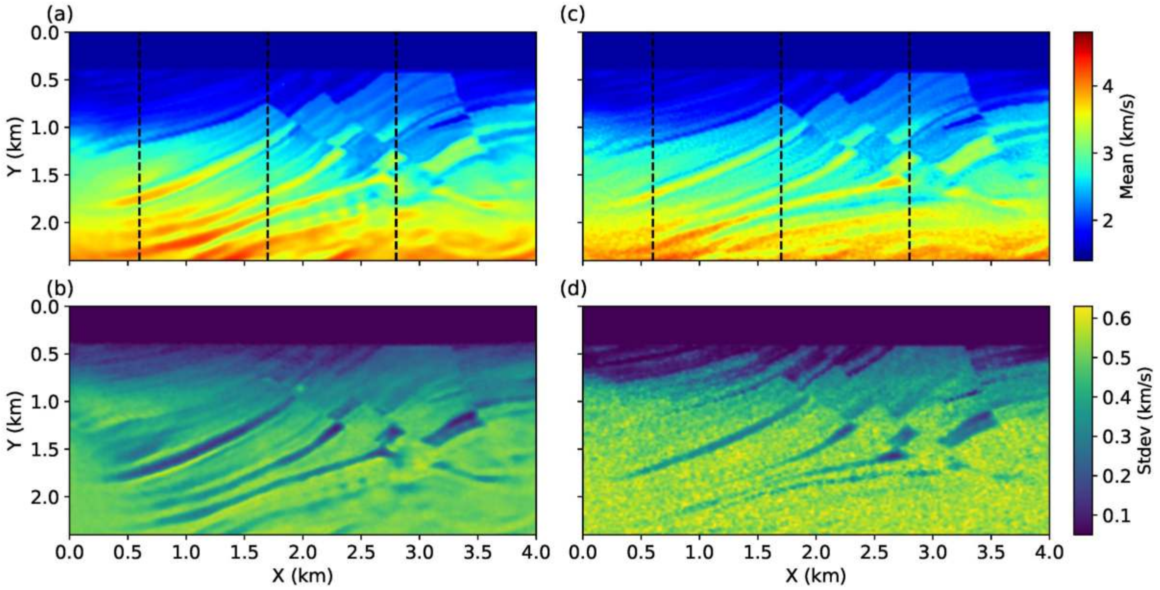


Figure 5 The mean (top row) and standard deviation (bottom row) obtained using SVGD (left panel) and sSVGD (right panel), respectively. Black dashed lines denote well log locations referred to in the main text.

Method	Number of simulations
SVGD	360,000
sSVGD	120,000

Table 2 Computational cost required by SVGD and sSVGD for FWI.

obtained using sSVGD are broader than those obtained using SVGD, which again demonstrates that SVGD can underestimate uncertainties. Note that in the results obtained using SVGD some true velocities lie outside the high probability area at large depths ($>$ 1.5 km), whereas those obtained using sSVGD generally include the true velocity in values with non-zero uncertainty. This shows that SVGD can produce biased results for high dimensional problems as noted in several studies (Ba et al., 2022; Zhang et al., 2023).

Similarly to the above section we measure the computational cost required by each method using the number of forward and adjoint simulations (Table 2). Specifically, SVGD required 360,000 simulations to converge, while sSVGD used 120,000 simulations. This again demonstrates that sSVGD can be more computationally efficient than SVGD because sSVGD requires fewer particles yet generates many more samples. To give an overall idea of the computational cost, the above inversions required 49 hours for sSVGD using 40 AMD EPYC CPU cores, and 3 days for SVGD using 90 CPU cores.

5 Discussion

Although in the VIP package we only implemented 2D travel time tomography and 2D full waveform inversion, the code can easily be applied to other types of problems, and also to larger scale problems by using modern high performance computation (HPC) facilities. For example, users can implement 3D full waveform inversion by providing a 3D forward and adjoint simulation code (see more details in the code documentation, and an example in Zhang et al., 2023). In order to enable easy deployment on HPC facilities, the code provides a guide on how to parallelize the computation using the Sun Grid Engine queuing system. Other queuing systems can be implemented in a similar way.

Although we have demonstrated that sSVGD can generate more accurate results than SVGD in high dimensional problems and requires less computational cost in total, the method generally requires many more iterations. As a result, sSVGD may be less efficient than SVGD in wall clock time when a large number of CPU cores is available. This is why we implement SVGD in the VIP package as in practice it may be a better choice for low dimensional problems.

ADVI may become inefficient in a high dimensional space because of the increased size of the covariance matrix. To enable applications in such cases, we also implement a diagonal covariance matrix, that is, a mean-field approximation (Kucukelbir et al., 2017). In SVGD and sSVGD besides the radial basis function kernel used in above examples, the package also implements diagonal matrix-valued kernel functions which are constructed by combining a positive definite diagonal matrix $\mathbf{Q}$ and the radial basis function (Wang et al., 2019; Zhang and Curtis, 2021). The elements of $\mathbf{Q}$ can be set as the inverse of the variance calculated across particles (Zhang and Curtis, 2021).

To promote reproducibility and show how to use the code, we included several examples along with the code which can be used to reproduce those results obtained in the above section. We encourage interested readers to begin with these examples to familiarize themselves with the code. Finally, we note that VIP is actively being developed and expanded, and contributions from the community are welcome.

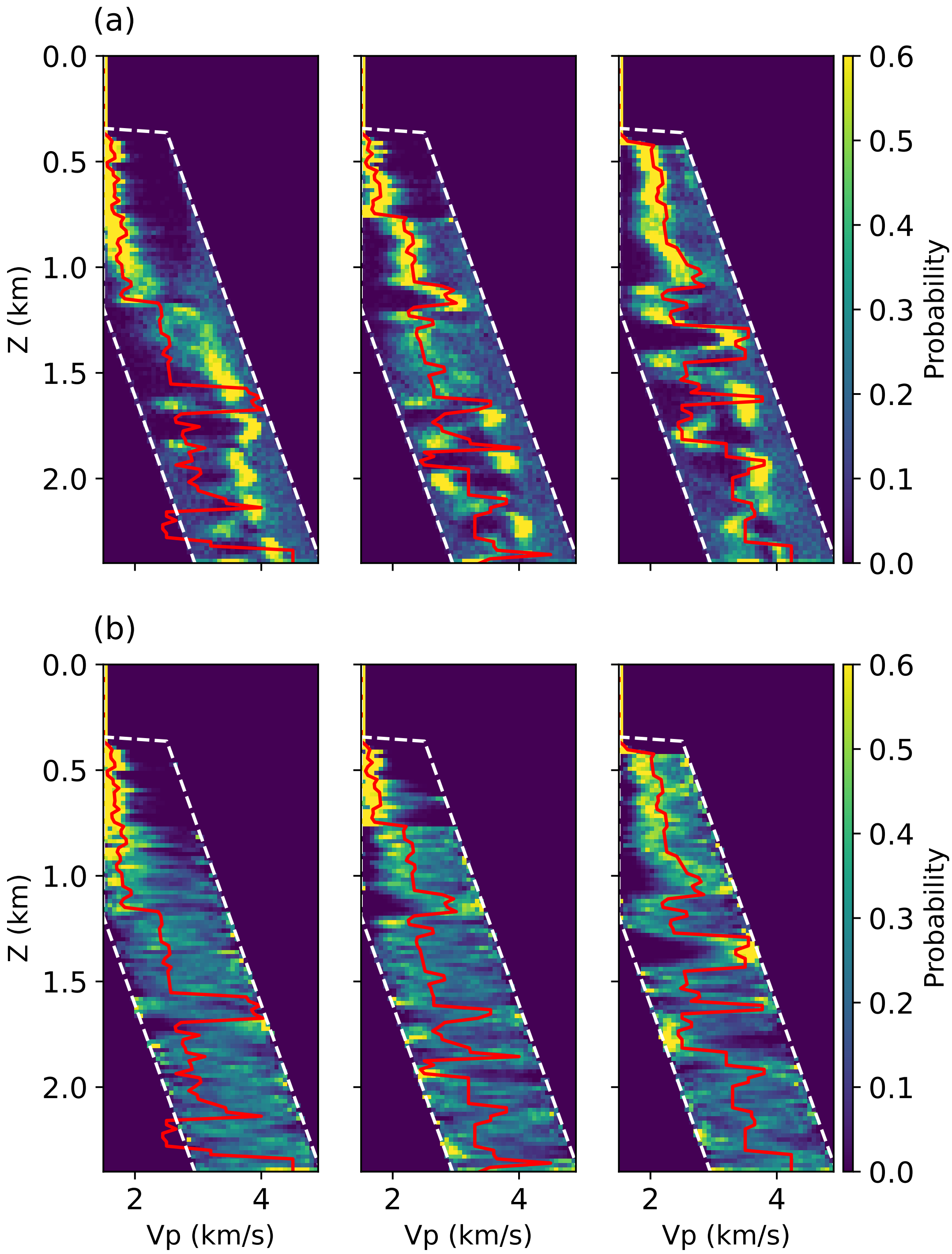


Figure 6 Marginal distributions at three well logs (black dashed lines in Figure 5) obtained using **(a)** SVGD and **(b)** sSVGD, respectively. Red lines show the true velocity profiles and white dashed lines show the lower and upper bound of the prior distribution.

6 Conclusion

VIP is a Python package which solves general inverse problems using variational inference methods, including automatic differential variational inference (ADVI), Stein variational gradient descent (SVGD) and stochastic SVGD (sSVGD). The package is designed to be easy enough for beginners to use, and efficient enough to solve complex inverse problems. In addition, VIP is implemented in a scalable way such that it can be deployed on a desktop as well as in high performance computation facilities. We demonstrated the package using two examples: 2D travel time tomography and 2D full waveform inversion. Users can also use the package to solve their own inverse problems by providing an appropriate forward modelling and gradient calculation code. We conclude that VIP can be used to solve a wide range of inverse problems in practice. The most recent release of the code can be downloaded from GitHub (https://github.com/xin2zhang/VIP) and a stable version is available on Zenodo (Zhang and Curtis, 2023).

Acknowledgements

The authors thank the Edinburgh Imaging Project sponsors (BP and Total), National Natural Science Foundation of China (42204055) and the Fundamental Research Funds for the Central Universities for supporting this research. This work has made use of the resources provided by the Edinburgh Compute and Data Facility (http://www.ecdf.ed.ac.uk/). The authors also thank two anonymous reviewers for their valuable feedback which significantly improved the manuscript.

Data and code availability

The code and data used in this study are available in a Zenodo repository (Zhang and Curtis, 2023).

References

Afanasiev, M., Boehm, C., van Driel, M., Krischer, L., Rietmann, M., May, D. A., Knepley, M. G., and Fichtner, A. Modular and flexible spectral-element waveform modelling in two and three dimensions. *Geophysical Journal International*, 216(3):1675–1692, 2019. doi: 10.1093/gji/ggy469.

Agata, R., Shiraishi, K., and Fujie, G. Bayesian seismic tomography based on velocity-space Stein variational gradient descent for physics-informed neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 2023. doi: 10.1109/TGRS.2023.3295414.

Ahmed, Z., Yunyue, L., and Arthur, C. Regularized seismic amplitude inversion via variational inference. *Geophysical Prospecting*, n/a(n/a), 2022. doi: 10.1111/1365-2478.13248.

Aki, K. and Lee, W. Determination of three-dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes: 1. A homogeneous initial model. *Journal of Geophysical research*, 81(23):4381–4399, 1976. doi: 10.1029/JB081i023p04381.

Andersen, K. E., Brooks, S. P., and Hansen, M. B. A Bayesian approach to crack detection in electrically conducting media. *Inverse Problems*, 17(1):121, 2001. doi: 10.1088/0266-5611/17/1/310.

Arnold, R. and Curtis, A. Interrogation theory. *Geophysical Journal International*, 214(3):1830–1846, 2018. doi: 10.1093/gji/ggy248.

Aster, R. C., Borchers, B., and Thurber, C. H. *Parameter estimation and inverse problems*. Elsevier, 2018.

Ba, J., Erdogdu, M. A., Ghassemi, M., Sun, S., Suzuki, T., Wu, D., and Zhang, T. Understanding the Variance Collapse of SVGD in High Dimensions. In *International Conference on Learning Representations*, 2022. https://openreview.net/forum?id=Qycd9j5Qp9J.

Behnel, S., Bradshaw, R., Citro, C., Dalcin, L., Seljebotn, D. S., and Smith, K. Cython: The best of both worlds. *Computing in Science & Engineering*, 13(2):31–39, 2010. doi: 10.1109/MCSE.2010.118.

Bishop, C. M. *Pattern recognition and machine learning*. springer, 2006.

Blei, D. M., Kucukelbir, A., and McAuliffe, J. D. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112(518):859–877, 2017. doi: 10.1080/01621459.2017.1285773.

Bodin, T. and Sambridge, M. Seismic tomography with the reversible jump algorithm. *Geophysical Journal International*, 178(3):1411–1436, 2009. doi: 10.1111/j.1365-246X.2009.04226.x.

Brooks, S., Gelman, A., Jones, G., and Meng, X.-L. *Handbook of Markov chain Monte Carlo*. CRC press, 2011.

Chen, T., Fox, E., and Guestrin, C. Stochastic Gradient Hamiltonian Monte Carlo. In Xing, E. P. and Jebara, T., editors, *Proceedings of the 31st International Conference on Machine Learning*, volume 32 of *Proceedings of Machine Learning Research*, pages 1683–1691, Bejing, China, 22–24 Jun 2014. PMLR. https://proceedings.mlr.press/v32/cheni14.html.

Curtis, A. and Lomax, A. Prior information, sampling distributions, and the curse of dimensionality. *Geophysics*, 66(2):372–378, 2001. doi: 10.1190/1.1444928.

Dosso, S. E., Holland, C. W., and Sambridge, M. Parallel tempering for strongly nonlinear geoacoustic inversion. *The Journal of the Acoustical Society of America*, 132(5):3030–3040, 2012. doi: 10.1121/1.4757639.

Duane, S., Kennedy, A. D., Pendleton, B. J., and Roweth, D. Hybrid Monte Carlo. *Physics letters B*, 195(2):216–222, 1987. doi: 10.1016/0370-2693(87)91197-X.

Duchi, J., Hazan, E., and Singer, Y. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12(Jul):2121–2159, 2011. http://jmlr.org/papers/v12/duchi11a.html.

Fichtner, A., Bunge, H.-P., and Igel, H. The adjoint method in seismology: I. Theory. *Physics of the Earth and Planetary Interiors*, 157(1-2):86–104, 2006. doi: 10.1016/j.pepi.2006.03.016.

Fichtner, A., Zunino, A., and Gebraad, L. Hamiltonian Monte Carlo solution of tomographic inverse problems. *Geophysical Journal International*, 216(2):1344–1363, 2018. doi: 10.1093/gji/ggy496.

Galetti, E., Curtis, A., Meles, G. A., and Baptie, B. Uncertainty loops in travel-time tomography from nonlinear wave physics. *Physical review letters*, 114(14):148501, 2015. doi: 10.1103/physrevlett.114.148501.

Galetti, E., Curtis, A., Baptie, B., Jenkins, D., and Nicolson, H. Transdimensional Love-wave tomography of the British Isles and shear-velocity structure of the East Irish Sea Basin from ambient-noise interferometry. *Geophysical Journal International*, 208(1):36–58, 2017. doi: 10.1093/gji/ggw286.

Gallagher, K., Charvin, K., Nielsen, S., Sambridge, M., and Stephenson, J. Markov chain Monte Carlo (MCMC) sampling methods to determine optimal models, model resolution and model choice for Earth Science problems. *Marine and Petroleum Geology*, 26(4):525–535, 2009. doi: 10.1016/j.marpetgeo.2009.01.003.

Gallego, V. and Insua, D. R. Stochastic gradient MCMC with repulsive forces. *arXiv preprint arXiv:1812.00071*, 2018.

Gebraad, L., Boehm, C., and Fichtner, A. Bayesian Elastic Full-Waveform Inversion Using Hamiltonian Monte Carlo. *Journal of Geophysical Research: Solid Earth*, 125(3):e2019JB018428, 2020. doi: 10.1029/2019JB018428.

Gong, C., Peng, J., and Liu, Q. Quantile Stein Variational Gradient Descent for Batch Bayesian Optimization. In Chaudhuri, K. and Salakhutdinov, R., editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 2347–2356. PMLR, 09–15 Jun 2019. https://proceedings.mlr.press/v97/gong19b.html.

Green, P. J. Reversible jump Markov chain Monte Carlo computation and Byesian model determination. *Biometrika*, pages 711–732, 1995. doi: 10.1093/biomet/82.4.711.

Hastings, W. K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970. doi: 10.1093/biomet/57.1.97.

Hawkins, R. and Sambridge, M. Geophysical imaging using transdimensional trees. *Geophysical Journal International*, 203(2):972–1000, 2015. doi: 10.1093/gji/ggv326.

Hukushima, K. and Nemoto, K. Exchange Monte Carlo method and application to spin glass simulations. *Journal of the Physical Society of Japan*, 65(6):1604–1608, 1996. doi: 10.1143/JPSJ.65.1604.

Izzatullah, M., Alkhalifah, T., Romero, J., Corrales, M., Luiken, N., and Ravasi, M. Plug-and-Play Stein variational gradient descent for Bayesian post-stack seismic inversion. In *84th EAGE Annual Conference & Exhibition*, volume 2023, pages 1–5. European Association of Geoscientists & Engineers, 2023. doi: 10.3997/2214-4609.202310177.

Kingma, D. P. and Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Komatitsch, D., Tromp, J., Garg, R., Gharti, H. N., Nagaso, M., Oral, E., and et al. SPECFEM/specfem3d: SPECFEM3D v4.1.0. doi: 10.5281/zenodo.10413988.

Kubrusly, C. and Gravier, J. Stochastic approximation algorithms and applications. In *1973 IEEE conference on decision and control including the 12th symposium on adaptive processes*, pages 763–766. IEEE, 1973. doi: 10.1109/CDC.1973.269114.

Kucukelbir, A., Tran, D., Ranganath, R., Gelman, A., and Blei, D. M. Automatic differentiation variational inference. *The Journal of Machine Learning Research*, 18(14):1–45, 2017. http://jmlr.org/papers/v18/16-107.html.

Kullback, S. and Leibler, R. A. On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86, 1951. http://www.jstor.org/stable/2236703.

Kumar, R., Kotsi, M., Siahkoohi, A., and Malcolm, A. Enabling uncertainty quantification for seismic data preprocessing using normalizing flows (NF)—An interpolation example. In *First International Meeting for Applied Geoscience & Energy*, pages 1515–1519. Society of Exploration Geophysicists, 2021. doi: 10.1190/segam2021-3583705.1.

Li, X., Bürgi, P. M., Ma, W., Noh, H. Y., Wald, D. J., and Xu, S. DisasterNet: Causal Bayesian Networks with Normalizing Flows for Cascading Hazards Estimation from Satellite Imagery. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 4391–4403, 2023. doi: 10.1145/3580305.3599807.

Lions, J. L. *Optimal control of systems governed by partial differential equations*, volume 170. Springer, 1971.

Liu, Q. and Wang, D. Stein Variational Gradient Descent: A General Purpose Bayesian Inference Algorithm. In Lee, D., Sugiyama, M., Luxburg, U., Guyon, I., and Garnett, R., editors, *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc., 2016. https://proceedings.neurips.cc/paper_files/paper/2016/file/b3ba8f1bee1238a2f37603d90b58898d-Paper.pdf.

Lomas, A., Luo, S., Irakarama, M., Johnston, R., Vyas, M., and Shen, X. 3D Probabilistic Full Waveform Inversion: Application to Gulf of Mexico Field Data. In *84th EAGE Annual Conference & Exhibition*, volume 2023, pages 1–5. European Association of Geoscientists & Engineers, 2023. doi: 10.3997/2214-4609.202310720.

Ma, Y.-A., Chen, T., and Fox, E. A Complete Recipe for Stochastic Gradient MCMC. 28, 2015. https://proceedings.neurips.cc/paper_files/paper/2015/file/9a4400501febb2a95e79248486a5f6d3-Paper.pdf.

Malinverno, A. Parsimonious Byesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem. *Geophysical Journal International*, 151(3):675–688, 2002. doi: 10.1046/j.1365-246X.2002.01847.x.

Malinverno, A., Leaney, S., et al. A Monte Carlo method to quantify uncertainty in the inversion of zero-offset VSP data. In *2000 SEG Annual Meeting*. Society of Exploration Geophysicists, 2000. doi: 10.1190/1.1815943.

Martin, G. S., Wiley, R., and Marfurt, K. J. Marmousi2: An elastic upgrade for Marmousi. *The leading edge*, 25(2):156–166, 2006. doi: 10.1190/1.2172306.

Martin, J., Wilcox, L. C., Burstedde, C., and Ghattas, O. A stochastic Newton MCMC method for large-scale statistical inverse problems with application to seismic inversion. *SIAM Journal on Scientific Computing*, 34(3):A1460–A1487, 2012. doi: 10.1137/110845598.

McKean, S., Priest, J., Dettmer, J., Fradelizio, G., and Eaton, D. Separating Hydraulic Fracturing Microseismicity From Induced Seismicity by Bayesian Inference of Non-Linear Pressure Diffusivity. *Geophysical Research Letters*, 50(14):e2022GL102131, 2023. doi: 10.1029/2022GL102131.

Metropolis, N. and Ulam, S. The Monte Carlo method. *Journal of the American statistical association*, 44(247):335–341, 1949. doi: 10.1080/01621459.1949.10483310.

Mosegaard, K. and Sambridge, M. Monte Carlo analysis of inverse problems. *Inverse problems*, 18(3):R29, 2002. doi: 10.1088/0266-5611/18/3/201.

Mosegaard, K. and Tarantola, A. Monte Carlo sampling of solutions to inverse problems. *Journal of Geophysical Research: Solid Earth*, 100(B7):12431–12447, 1995. doi: 10.1029/94JB03097.

Nawaz, M. A. and Curtis, A. Variational Bayesian inversion (VBI) of quasi-localized seismic attributes for the spatial distribution of geological facies. *Geophysical Journal International*, 214(2):845–875, 2018. doi: 10.1093/gji/ggy163.

Nawaz, M. A. and Curtis, A. Rapid Discriminative Variational Byesian Inversion of Geophysical Data for the Spatial Distribution of Geological Properties. *Journal of Geophysical Research: Solid Earth*, 2019. doi: 10.1029/2018JB016652.

Nawaz, M. A., Curtis, A., Shahraeeni, M. S., and Gerea, C. VARIATIONAL BAYESIAN INVERSION OF SEISMIC ATTRIBUTES JOINTLY FOR GEOLOGICAL FACIES AND PETROPHYSICAL ROCK PROPERTIES. *GEOPHYSICS*, pages 1–78, 2020. doi: 10.1190/geo2019-0163.1.

Nicolson, H., Curtis, A., Baptie, B., and Galetti, E. Seismic interferometry and ambient noise tomography in the British Isles. *Proceedings of the Geologists' Association*, 123(1):74–86, 2012. doi: 10.1016/j.pgeola.2011.04.002.

Nicolson, H., Curtis, A., and Baptie, B. Rayleigh wave tomography of the British Isles from ambient seismic noise. *Geophysical Journal International*, 198(2):637–655, 2014. doi: 10.1093/gji/ggu071.

O'Hagan, A. and Forster, J. J. *Kendall's advanced theory of statistics, volume 2B: Bayesian inference*, volume 2. Arnold, 2004.

Oksendal, B. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.

Orozco, R., Louboutin, M., Siahkoohi, A., Rizzuti, G., van Leeuwen, T., and Herrmann, F. Amortized Normalizing Flows for Transcranial Ultrasound with Uncertainty Quantification. *arXiv preprint arXiv:2303.03478*, 2023. doi: 10.48550/arXiv.2303.03478.

Parisi, G. *Statistical field theory*. Addison-Wesley, 1988. doi: 10.1063/1.2811677.

Pinder, T., Nemeth, C., and Leslie, D. Stein variational Gaussian processes. *arXiv preprint arXiv:2009.12141*, 2020. doi: 10.48550/arXiv.2009.12141.

Plessix, R.-E. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167(2):495–503, 2006. doi: 10.1111/j.1365-246X.2006.02978.x.

Ramgraber, M., Weatherl, R., Blumensaat, F., and Schirmer, M. Non-Gaussian Parameter Inference for Hydrogeological Models Using Stein Variational Gradient Descent. *Wa-*

ter Resources Research, 57(4):e2020WR029339, 2021. doi: 10.1029/2020WR029339.

Ramirez, A. L., Nitao, J. J., Hanley, W. G., Aines, R., Glaser, R. E., Sengupta, S. K., Dyer, K. M., Hickling, T. L., and Daily, W. D. Stochastic inversion of electrical resistivity changes using a Markov Chain Monte Carlo approach. *Journal of Geophysical Research: Solid Earth*, 110(B2), 2005. doi: 10.1029/2004JB003449.

Rawlinson, N. FMST: fast marching surface tomography package–Instructions. *Research School of Earth Sciences, Australian National University, Canberra*, 29:47, 2005.

Rawlinson, N. and Sambridge, M. Multiple reflection and transmission phases in complex layered media using a multistage fast marching method. *Geophysics*, 69(5):1338–1350, 2004. doi: 10.1190/1.1801950.

Rezende, D. J. and Mohamed, S. Variational inference with normalizing flows. *arXiv preprint arXiv:1505.05770*, 2015. doi: 10.48550/arxiv.1505.05770.

Robbins, H. and Monro, S. A stochastic approximation method. *The annals of mathematical statistics*, pages 400–407, 1951. doi: 10.1214/aoms/1177729586.

Roberts, G. O., Tweedie, R. L., et al. Exponential convergence of Langevin distributions and their discrete approximations. *Bernoulli*, 2(4):341–363, 1996. doi: 10.2307/3318418.

Rocklin, M. et al. Dask: Parallel computation with blocked algorithms and task scheduling. In *Proceedings of the 14th python in science conference*, volume 130, page 136. SciPy Austin, TX, 2015. doi: 10.25080/majora-7b98e3ed-013.

Rücker, C., Günther, T., and Wagner, F. M. pyGIMLi: An open-source library for modelling and inversion in geophysics. *Computers & Geosciences*, 109:106–123, 2017. doi: 10.1016/j.cageo.2017.07.011.

Sambridge, M. A parallel tempering algorithm for probabilistic sampling and multimodal optimization. *Geophysical Journal International*, page ggt342, 2013. doi: 10.1093/gji/ggt342.

Sambridge, M. and Mosegaard, K. Monte Carlo methods in geophysical inverse problems. *Reviews of Geophysics*, 40(3):3–1, 2002. doi: 10.1029/2000RG000089.

Sen, M. K. and Biswas, R. Transdimensional seismic inversion using the reversible jump Hamiltonian Monte Carlo algorithm. *Geophysics*, 82(3):R119–R134, 2017. doi: 10.1190/geo2016-0010.1.

Shen, W., Ritzwoller, M. H., Schulte-Pelkum, V., and Lin, F.-C. Joint inversion of surface wave dispersion and receiver functions: a Byesian Monte-Carlo approach. *Geophysical Journal International*, 192(2):807–836, 2012. doi: 10.1093/gji/ggs050.

Siahkoohi, A., Rizzuti, G., Witte, P. A., and Herrmann, F. J. Faster Uncertainty Quantification for Inverse Problems with Conditional Normalizing Flows. *arXiv preprint arXiv:2007.07985*, 2020. doi: 10.48550/arXiv.2007.07985.

Siahkoohi, A., Orozco, R., Rizzuti, G., and Herrmann, F. J. Wave-equation-based inversion with amortized variational Bayesian inference. *arXiv preprint arXiv:2203.15881*, 2022a. doi: 10.48550/arXiv.2203.15881.

Siahkoohi, A., Rizzuti, G., and Herrmann, F. J. Deep Bayesian inference for seismic imaging with tasks. *Geophysics*, 87(5): S281–S302, 2022b. doi: 10.1190/geo2021-0666.1.

Smith, J. D., Ross, Z. E., Azizzadenesheli, K., and Muir, J. B. HypoSVI: Hypocentre inversion with Stein variational inference and physics informed neural networks. *Geophysical Journal International*, 228(1):698–710, 2022. doi: 10.1093/gji/ggab309.

Tarantola, A. Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8):1259–1266, 1984. doi: 10.1190/1.1441754.

Tarantola, A. Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation. In *Scattering and Attenuations of Seismic Waves, Part I*, pages 365–399. Springer, 1988. doi: 10.1007/bf01772605.

Tarantola, A. *Inverse problem theory and methods for model parameter estimation*, volume 89. SIAM, 2005. doi: 10.1137/1.9780898717921.

Team, S. D. et al. Stan modeling language users guide and reference manual. *Technical report*, 2016.

Tromp, J., Tape, C., and Liu, Q. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International*, 160(1):195–216, 2005. doi: 10.1111/j.1365-246X.2004.02453.x.

Valentine, A. P. and Sambridge, M. Gaussian process models—I. A framework for probabilistic continuous inverse theory. *Geophysical Journal International*, 220(3):1632–1647, 2020. doi: 10.1093/gji/ggz520.

Wang, D., Tang, Z., Bajaj, C., and Liu, Q. Stein variational gradient descent with matrix-valued kernels. In *Advances in neural information processing systems*, pages 7836–7846, 2019. doi: 10.48550/arXiv.1910.12794.

Wang, W., McMechan, G. A., and Ma, J. Re-weighted variational full waveform inversions. *Geophysics*, 88(4):1–61, 2023. doi: 10.1190/geo2021-0766.1.

Wathelet, M., Chatelain, J.-L., Cornou, C., Giulio, G. D., Guillier, B., Ohrnberger, M., and Savvaidis, A. Geopsy: A user-friendly open-source tool set for ambient vibration processing. *Seismological Research Letters*, 91(3):1878–1889, 2020. doi: 10.1785/0220190360.

Welling, M. and Teh, Y. W. Bayesian learning via stochastic gradient Langevin dynamics. In *Proceedings of the 28th international conference on machine learning (ICML-11)*, pages 681–688. Citeseer, 2011.

Zaroli, C. Global seismic tomography using Backus–Gilbert inversion. *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, 207(2):876–888, 2016. doi: 10.1093/gji/ggw315.

Zaroli, C. Seismic tomography using parameter-free Backus–Gilbert inversion. *Geophysical Journal International*, 218(1): 619–630, 2019. doi: 10.1093/gji/ggz175.

Zaroli, C., Koelemeijer, P., and Lambotte, S. Toward seeing the Earth's interior through unbiased tomographic lenses. *Geophysical Research Letters*, 44(22):11–399, 2017. doi: 10.1002/2017GL074996.

Zeiler, M. D. Adadelta: an adaptive learning rate method. *arXiv preprint arXiv:1212.5701*, 2012. doi: 10.48550/arXiv.1212.5701.

Zhang, C. and Chen, T. Bayesian slip inversion with automatic differentiation variational inference. *Geophysical Journal International*, 229(1):546–565, 2022. doi: 10.1093/gji/ggab438.

Zhang, C., Bütepage, J., Kjellström, H., and Mandt, S. Advances in variational inference. *IEEE transactions on pattern analysis and machine intelligence*, 41(8):2008–2026, 2018a. doi: 10.1109/TPAMI.2018.2889774.

Zhang, X. and Curtis, A. Seismic tomography using variational inference methods. *Journal of Geophysical Research: Solid Earth*, 125(4):e2019JB018589, 2020a.

Zhang, X. and Curtis, A. Variational full-waveform inversion. *Geophysical Journal International*, 222(1):406–411, 2020b. doi: 10.1093/gji/ggaa170.

Zhang, X. and Curtis, A. Bayesian Full-waveform Inversion with Realistic Priors. *Geophysics*, 86(5):1–20, 2021. doi: 10.1190/geo2021-0118.1.

Zhang, X. and Curtis, A. Interrogating probabilistic inversion results for subsurface structural information. *Geophysical Journal International*, 229(2):750–757, 2022. doi: 10.1093/gji/ggab496.

Zhang, X., Curtis, A., Galetti, E., and de Ridder, S. 3-D Monte Carlo surface wave tomography. *Geophysical Journal International*, 215(3):1644–1658, 2018b. doi: 10.1093/gji/ggy362.

Zhang, X., Lomas, A., Zhou, M., Zheng, Y., and Curtis, A. 3-D Bayesian variational full waveform inversion. *Geophysical Journal International*, 234(1):546–561, 2023. doi: 10.1093/gji/ggad057.

Zhang, Z. and Curtis, A. Variational Inversion Package, 2023. doi: 10.5281/zenodo.10036815.

Zhao, X. and Curtis, A. Bayesian inversion, uncertainty analysis and interrogation using boosting variational inference. *Journal of Geophysical Research: Solid Earth*, 129(1):e2023JB027789, 2024. doi: 10.1029/2023JB027789.

Zhao, X., Curtis, A., and Zhang, X. Bayesian seismic tomography using normalizing flows. *Geophysical Journal International*, 228 (1):213–239, 2022a.

Zhao, X., Curtis, A., and Zhang, X. Interrogating Subsurface Structures using Probabilistic Tomography: an example assessing the volume of Irish Sea basins. *Journal of Geophysical Research: Solid Earth*, 127(4):e2022JB024098, 2022b. doi: 10.1029/2022JB024098.

Zhao, Z. and Sen, M. K. A gradient based MCMC method for FWI and uncertainty analysis. In *SEG Technical Program Expanded Abstracts 2019*, pages 1465–1469. Society of Exploration Geophysicists, 2019. doi: 10.1190/segam2019-3216560.1.

Zhdanov, M. S. *Geophysical inverse theory and regularization problems*, volume 36. Elsevier, 2002. doi: 10.1016/s0076-6895(02)x8037-8.

Zunino, A., Gebraad, L., Ghirotto, A., and Fichtner, A. HMCLab: a framework for solving diverse geophysical inverse problems using the Hamiltonian Monte Carlo method. *arXiv preprint arXiv:2303.10047*, 2023. doi: 10.1093/gji/ggad403.

REPORT
doi:10.26443/seismica.v3i1.1099

SeisMIC - an Open Source Python Toolset to Compute Velocity Changes from Ambient Seismic Noise

Peter Makus * [1,2], **Christoph Sens-Schönfelder** [1]

[1]Helmholtz Center, German Research Center for Geosciences GFZ, Potsdam, Germany, [2]Institute for Geological Sciences, Freie Universität Berlin, Berlin, Germany

Author contributions: *Conceptualization*: PM, CS. *Methodology*: CS, PM. *Software*: PM, CS. *Formal Analysis*: PM. *Writing - Original draft*: PM. *Writing - Review & Editing*: CS. *Supervision*: CS. *Funding acquisition*: CS.

Abstract We present SeisMIC, a fast, versatile, and adaptable open-source software to estimate seismic velocity changes from ambient seismic noise. SeisMIC includes a broad set of tools and functions to facilitate end-to-end processing of ambient noise data, from data retrieval and raw data analysis via spectrogram computation, over waveform coherence analysis, to post-processing of the final velocity change estimates. A particular highlight of the software is its ability to invert velocity change time series onto a spatial grid, making it possible to create maps of velocity changes. With the software, we implement new data formats ensuring uniformity, flexibility, interoperability, and integrity. To tackle the challenge of processing large continuous datasets, SeisMIC can exploit multithreading at high efficiency with an about five-time improvement in compute time compared to MSNoise, probably the most widespread ambient noise software. In this manuscript, we provide a short tutorial and tips for users on how to employ SeisMIC most effectively. Extensive and up-to-date documentation is available online. Its broad functionality combined with easy adaptability and high efficiency make SeisMIC a well-suited tool for studies across all scales.

Production Editor:
Gareth Funning
Handling Editor:
Paula Koelemeijer
Copy & Layout Editor:
Théa Ragon

Signed reviewer(s):
Laura Ermert

Received:
August 25, 2023
Accepted:
January 24, 2024
Published:
February 2, 2024

1 Introduction

Over the past twenty years, the analysis of temporal changes in seismic velocity has become a standard tool in seismology. Seismologists exploit records of repeating sources, such as explosives (e.g., Nishimura et al., 2000; Hirose et al., 2017), vibrators (e.g., Clymer and McEvilly, 1981; Ikuta et al., 2002), airguns (e.g., Wegler et al., 2006; Yang et al., 2018), or earthquake doublets (e.g., Poupinet et al., 1984; Sawazaki et al., 2015), to quantify such changes. Commonly, the analysis of delays focuses on the later arriving, multiply scattered wave train - the so-called coda, which samples the medium to a greater spatial extent than the first-arriving energy and is sensitive even to minute velocity changes (dv/v) in the order of per-mills (Snieder et al., 2002). We refer to this technique as coda wave interferometry.

While active source coda wave interferometry accurately resolves dv/v, studies using artificial sources are logistically challenging and expensive. Repeating natural sources, on the other hand, rarely occur in regular patterns, allowing only for a coarse temporal resolution of dv/v in seismically active regions. Sens-Schönfelder and Wegler (2006) obtained dv/v by analysing modifications in the correlations of continuous waveforms. Their method, passive image interferometry (PII), relies on the diffusive energy field of the ubiquitous ambient seismic noise (Sens-Schönfelder and Wegler, 2011). PII has successfully been applied to quantify velocity changes, for example due to seasonal meteorological cycles (e.g., Sens-Schönfelder and Wegler, 2006; Wang et al., 2017), earthquake damage (e.g., Brenguier et al., 2008; Minato et al., 2012), volcanic deformation (e.g., Sens-Schönfelder et al., 2014b; Donaldson et al., 2019), groundwater fluctuations (e.g., Clements and Denolle, 2018; Illien et al., 2021; Mao et al., 2022), landslides (e.g., Bièvre et al., 2018), or climate-change-induced thawing (e.g., Mordret et al., 2016; Lindner et al., 2021). This breadth of applications makes PII a widely used methodology.

Processing and analysing continuous waveforms comes with multiple challenges due to the large amount of raw and derived data, such as the need for efficient processing and storage strategies (Arrowsmith et al., 2022). Still today, many authors use unpublished codes to produce results for later publication and interpretation making it difficult for fellow researchers to reproduce or adapt the analyses. Using community codes published in the spirit of the FAIR principles (Barker et al., 2022) can facilitate the reproducibility of research, exchange in the community, and progress in science. Only a few software solutions exist for ambient noise seismology. Perhaps the most popular among these are MSNoise (Lecocq et al., 2014) and NoisePy (Jiang and Denolle, 2020). However, as we will show and discuss here, the existing software still leaves a niche to fill. For example, MSNoise is more specialised for end-to-end workflows and automated monitoring solutions, lending it more towards applications in large observatories, whereas, recently, NoisePy has undergone development towards cloud computing. To fill the remain-

*Corresponding author: makus@gfz-potsdam.de

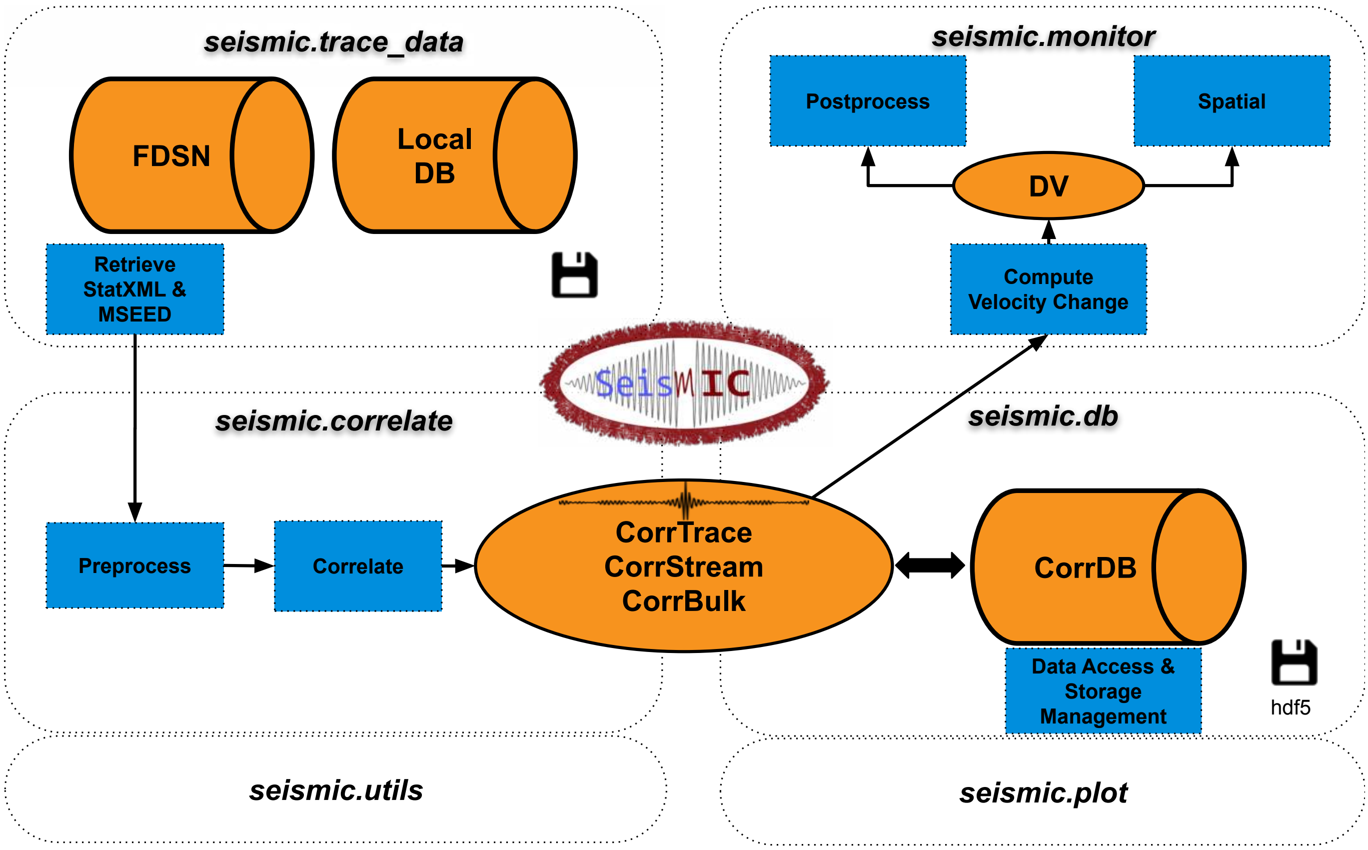


Figure 1 A flowchart summarising SeisMIC's modules and their purposes. A general workflow starts with data retrieval, continues with the computation of correlation functions, from which a velocity change time series can subsequently be estimated. We illustrate this with the example given in section 3. The depicted floppy disk marks database management modules. Operations and processes are shown in blue, whereas objects and databases are shown in orange. For the sake of simplicity, we omit non-essential objects and functions, instead, the flowchart focuses on the core processes.

ing gap, we introduce SeisMIC (Seismological Monitoring using Interferometric Concepts, Makus and Sens-Schönfelder, 2022), a fast, robust, flexible, and easily adapted Python tool to compute, process, and analyse dv/v. Due to these attributes, SeisMIC especially excels in the analysis of campaign data, where both ease of use and flexibility are crucial.

2 Modular Structure

2.1 Whom is it for? - The Philosophy behind SeisMIC

As outlined above, monitoring surveys are applied to a broad spectrum of research scopes resulting in a high diversity of requirements for research software. With that in mind, we developed SeisMIC to be flexible and adaptable to user needs. As opposed to working with a black box, users work close to the source code, making it easy to develop individualised workflows. Modules, submodules, or even single objects and functions of the code can also be used individually. Yet, the software remains a light and fast package, in which we avoid overhead due to non-essential functionality. For example, in contrast to MSNoise, we avoid heavy database management structure for continuous observatory monitoring, resulting in a significantly faster processing (see section 2.3.2) and giving SeisMIC an advantage in the analysis of campaign based data.

Learning to use a new code and even only determining whether a code satisfies one's need is a large time investment. To guarantee a fast start and a steep learning curve, we aligned SeisMIC closely with ObsPy (Beyreuther et al., 2010), with whose syntax almost all seismologists are familiar. In addition, we host tutorials and extensive, regularly-updated documentation at https://petermakus.github.io/SeisMIC/. All objects, methods, and functions have documentation strings according to the Sphinx standard.

As developers, we follow the FAIR principles (Hong et al., 2022). That is, we make SeisMIC findable, accessible, interoperable, and reusable. SeisMIC is a community code with clearly communicated community standards, and users can discuss or report issues, suggest changes, or submit pull requests via GitHub. We distribute SeisMIC under the European Union Public License 1.2.

Lastly, we keep up to high standards regarding functional robustness. We test functional integrity using a combination of integral and unit tests. To date, SeisMIC has successfully been applied to a broad range of applications, such as volcanic environments (Makus et al., 2023b,a), lab-scale applications (Asnar et al., 2023), and cryoseismological analyses (Nanni et al., 2023).

2.2 Implementation

As commonplace in Python, we structure SeisMIC in a modular fashion. We divide the program into clear modules, which, in turn, are subdivided into submodules. These modules can either be used separately or connected into a workflow/pipeline, starting from data retrieval and concluding with the computation, plotting, and postprocessing of dv/v objects. We show a chart with a simplified overview of SeisMIC's modular structure in Figure 1.

As shown in Figure 1, SeisMIC consists of four main modules. `seismic.trace_data` hosts the code for reading raw waveform data and station information. Alternatively, it can request data from FDSN servers. SeisMIC handles waveform data in *miniseed* format in daily chunks, while it saves station information in *StationXML* format. Generally, station response information is only necessary if the user opts to remove the station response before correlating. However, basic station information, such as the station's geographic coordinates, is always required.

All objects and functions to preprocess waveform data and compute correlation functions (CFs) are located in `seismic.correlate`. We include commonly used preprocessing functions such as detrending, tapering, amplitude clipping, sign-bit-normalisation, or spectral whitening (Bensen et al., 2007). For a complete and up-to-date list of preprocessing functions, consult SeisMIC's documentation. Users can easily import custom processing functions into the workflow. We compute CFs by transferring traces to matrices, computing the Fourier transform, and then computing their cross-correlation in the frequency domain. Suppose we want to calculate all available correlations from a dataset of M waveforms, of which each has N samples (indices m and n, respectively). Then, the respective mathematical operations can be expressed as follows:

First, we compute the discrete Fourier transform of the matrix s containing the waveforms in the time domain:

$$S_{m,k} = \sum_{n=1}^{N} s_{m,n} e^{-\frac{i2\pi}{N}kn} \tag{1}$$

where $i = \sqrt{-1}$ and k is the sample index of the signal in the frequency domain. Secondly, we obtain the correlation matrix $\mathbf{C}$ by computing the product of the matrix with the complex conjugate of itself. We then repeat the operation M times, each time rolling the complex conjugate matrix by $j = \{1, 2, .., M\}$ lines:

$$C_{o,k} = S_{m,k}\overline{S_{m+j,k}} \tag{2}$$

where the bar indicates the complex conjugate and o indexes the station pair. In the described scenario, we obtain M^2 CFs, which are subsequently transferred back to the time domain:

$$C_{o,n} = \frac{1}{N}\sum_{k=1}^{N} C_{o,k} e^{\frac{i2\pi}{N}kn} \tag{3}$$

The CFs are then stored as special objects with attributes, plotting and post-processing methods. Finally, SeisMIC writes the CFs to a storage- and computationally-efficient *HDF5* container (Koranne, 2011).

All functionality to estimate velocity changes from the CFs resides in `seismic.monitor`. Currently, SeisMIC supports the estimation of velocity changes using the stretching technique (Sens-Schönfelder and Wegler, 2006) and we are implementing the wavelet-cross-spectrum analysis (Mao et al., 2020).

The stretching technique compares a reference correlation function $\tilde{C}_n$ to a CF C_n^l computed from data at an arbitrary subwindow l of the total time series. Note that we omit the index o indicating the station pair since this operation is independently executed for each station pair. There are several approaches to obtaining $\tilde{\mathbf{C}}$, all with their unique advantages, SeisMIC supports the use of single or multiple references (Sens-Schönfelder et al., 2014b). In SeisMIC, we implemented a grid search, in which we evaluate $\tilde{\mathbf{C}}$ at a new time vector $\tilde{\tau}$ stretched (or compressed) with the stretching factor κ_j:

$$\tilde{\tau}_j = \tau e^{-\kappa_j} \tag{4}$$

Note that we base the exponential stretching on a Taylor extension for small velocity changes. This assumption is more accurate than the more common $\tilde{\tau}_j \approx \tau(1 + \kappa_j)$ and has the advantage of yielding linearly reversible stretched functions. In the supplementary material, we provide a derivation.

Using our stretched time vector, we obtain a stretched reference correlation matrix with J lines, where J is the total number of tested stretch factors. Afterwards, we compute the zero-lag correlation (i.e., the normalised dot product) between each stretched reference and $\mathbf{C}^l$:

$$R_j^l = \sum_{n=1}^{N} \tilde{C}_n^j C_n^l \left(\sum_{n=1}^{N} (\tilde{C}_n^j)^2 \sum_{n=1}^{N} (C_n^l)^2 \right)^{-1/2} \tag{5}$$

The stretching factor $\kappa_j = -dv/v$ resulting in the maximum R_j^l corresponds to the negative apparent velocity change at time step l. The maximum value of R measures the velocity change estimate's stability and is often referred to as coherence. We then compute R_j^l for all time steps resulting in the similarity matrix $\mathbf{R}$, the final velocity change time series, and a corresponding coherence time series. Note that $\mathbf{R}$ is usually not computed for the whole coda, but just for a user-defined subset of lag time samples. In SeisMIC, dv/v can either be jointly inverted from causal (right) and acausal (left side) or estimated from either side, which might be desirable for active source experiments or if one side of the CF exhibits a superior signal-to-noise-ratio.

Finally, the computed velocity change time series can be post-processed and plotted using pre-implemented or custom functions. In addition, SeisMIC can invert a set of velocity change time series from different stations onto a map using the inversion method described by Obermann et al. (2013). To our knowledge, SeisMIC is currently the only publicly available software that supports spatial inversion of velocity change time series.

The workflow steps outlined above rely entirely on well-known Python libraries, including *NumPy* (Harris et al., 2020), *SciPy* (Virtanen et al., 2020), *ObsPy*

Table 1 Extraction from the header of a correlation function computed in section 3.

Field name	Value	Explanation
network	X9-X9	SEED network codes, dash-separated
station	IR1-IR1	SEED station codes, dash-separated
channel	HHE-HHE	SEED channel codes, dash-separated
location	-	SEED location codes, dash-separated (may be empty)
corr_start	2016-01-25T01...	UTC start time of the correlated traces
corr_end	2016-02-25T01...	UTC end time of the correlated traces
start_lag	-25.0	computed start lag in seconds
⋮	⋮	⋮

(Beyreuther et al., 2010), *Matplotlib* (Hunter, 2007), and *h5py* (Collette et al., 2020). To ensure the best stability, we only utilise the most well-maintained projects and keep the number of dependencies to a minimum. Some of SeisMIC's core functionalities are based on the MIIC software project (Sens-Schönfelder et al., 2014a). SeisMIC's latest beta version 0.5.3 is compatible with Python 3.10 and 3.11.

2.2.1 Data Formats and Standards

At the time of writing, there are no established standards for data handling in ambient noise seismology that would facilitate the exchange of correlation functions and subsequent processing with different tools. In the seismological community, excellent examples of well-designed data representations that developed into quasi-standards are the ObsPy (Beyreuther et al., 2010) trace and stream classes for waveform data and the inventories for station metadata. Such successful representations require some core attributes:

1. Uniformity: Various datasets have the same set of attributes, making them directly comparable.
2. Easy and flexible I/O (i.e., input/output), where data can be read, modified and stored later. Reading and writing operations are fast and easy. Modifications can be stored safely.
3. Interoperability: Data can easily be imported and exported into broadly used applications or libraries, facilitating data exchange.
4. Integrity: The data format must contain all information required for later processing, analysis, or cataloguing. No crucial information should be lost.

With SeisMIC, we suggest a representation of noise correlation functions implementing these attributes. For correlation functions, we base our data representation on the successful ObsPy streams and traces by introducing the `CorrTrace` and `CorrStream` classes that incorporate the specific requirements of CFs to ensure uniformity and integrity.

For the storage of CFs, the seismological standard for waveform data, MiniSEED, is not appropriate since it does not allow for the storage of the required meta information. The solution provided in SeisMIC stores the data itself in the form of a NumPy array complemented with a header containing information about the recording and correlation computation, such as sample rate, start and duration of the correlated time windows, minimum and maximum lag times, seed identifiers of the used stations, and coordinates of these stations. We show an extract of the header fields for an exemplary dataset in Table 1. `CorrTrace` headers also contain information about executed processing steps, such as filtering or tapering. The naming of stations follows the SEED convention. To ensure interoperability, data and header can easily be converted into NumPy arrays and Python dictionaries, respectively. The objects come with processing and plotting methods. As outlined above, SeisMIC saves `CorrStreams` in hdf5 containers, from which they can later be read, modified, and saved again.

2.3 Benchmark and Performance

In ambient noise seismology, it is not uncommon to work with data volumes in the order of terabytes. We address the arising computational and storage challenges with efficient and high-performance computing (HPC) compatible code design. To this end, SeisMIC enables parallel computing of correlations, velocity change estimates and spatial inversions, where the computation of CFs is the most expensive operation by a large margin. We implement parallel computing using *mpi4py* (Dalcin and Fang, 2021), which relies on the message passing interface (MPI). In contrast to other Python multithreading solutions, MPI-based solutions work seamlessly on high-performance computing (HPC) and cluster solutions.

In SeisMIC, the computationally most expensive parts of the workflow described in section 2.2 are the calculation of correlation functions, the associated preprocessing, and the estimation of the final velocity change time series. Therefore, an effective parallelisation scheme matters the most in these steps. For users, it is also important to understand how memory requirements scale. For the computation of CFs and the preprocessing of raw data, each core reads different raw data in chunks of equal length (see Listing 3 for details). Subsequently, the same core performs the preprocessing. For the cross-correlation operation, each core is responsible for a different component combination. This implementation makes the RAM usage practically independent of the number of cores used. Thus, RAM usage will mainly depend on the length of the raw data

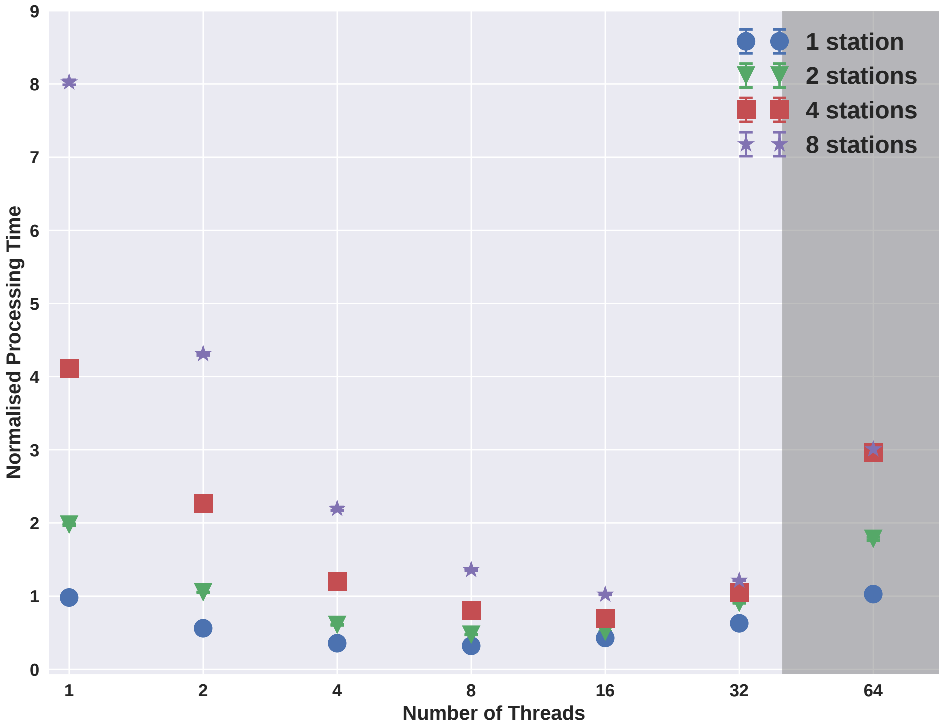


Figure 2 Multi-core scaling properties of SeisMIC. We show compute times for auto-correlations as a function of number of three-component datasets and number of parallel processing threads. The data points correspond to the mean processing time and the error bars to its standard deviation for ten operations (mostly too small to be visible). The processing times are normalised by the time needed to compute the correlations for one station using only one thread. The shaded area marks the area where the number of threads exceeds the number of physical cores, 40, i.e., the area where hyperthreading is employed.

chunks read in each step (i.e., a smaller read length will lead to lower memory usage) and its sampling rate (i.e., a lower sampling rate will lead to lower memory usage). Resulting CFs are written to h5 files immediately after correlation or stacking and the memory is freed. In contrast, SeisMIC computes the final dv/v estimate with ”1-core per component combination”. Here, a single core loads all available CFs for one component combination and executes the stretching algorithm and the associated processing. Therefore, for the final dv/v calculation, the memory requirement scales with the number of employed cores.

2.3.1 Multicore Scaling

To test how SeisMIC's computational performance scales with the number of used threads, we compute autocorrelations from three component data on a single cluster node featuring an Intel Cascadelake CPU structure that is equipped with 2 CPU sockets, each holding 20 physical cores that can each execute two threads in parallel. For our test, we compute CFs from 30 days of waveform data. SeisMIC reads daily chunks of miniseed files, which it subsequently decimates, here to a sampling rate of 25 Hz, after imposing an anti-alias filter. The daily waveforms are then detrended, tapered, and filtered with a pass band between 0.01 and 12 Hz. The data is then sliced into hourly traces, which are again linearly detrended, filtered between 2 and 8 Hz, and clipped if the amplitude exceeds a threshold of 2.5 times its standard deviation. Then, SeisMIC computes hourly CFs in the frequency domain and saves them in a customised HDF5 container after performing an inverse Fourier transform. We provide the *YAML* file containing the processing parameters in the supplementary material. We execute this operation using 1, 2, 4, 8, 16, 32, and 64 threads for data from 1, 2, 4, and 8 stations (i.e., 3, 6, 12, and 24 channels and component combinations). For each configuration, we repeat the computation ten times.

Figure 2 shows the mean processing time and standard deviation over the ten operations per unique $n_{threads}$-$n_{stations}$-combination. We normalise the processing times by the time required for $n_{threads} = 1$ and $n_{stations} = 1$. While $n_{threads} \leq n_{channels}$, where, in our case, $n_{channels} = 3n_{stations}$, the processing time scales close to linearly with the number of used threads, indicating an excellent parallel computing performance. As most of the parallel processing in SeisMIC works on a one-core-per-channel basis, only very little increase can be expected beyond this threshold. Indeed, for $n_{channels} < n_{threads}$, the code reaches a performance plateau. From here on, the processing time increases with a further increase of $n_{threads}$, probably due to MPI's communication overhead. Based on the shown results, we would discourage hyperthreading (i.e., using more threads than available physical cores), which leads to a significant performance drop. Generally, one should not employ more threads than the total number of available channels for the computation of correlation functions or the total number of channel combinations for the dv/v estimation.

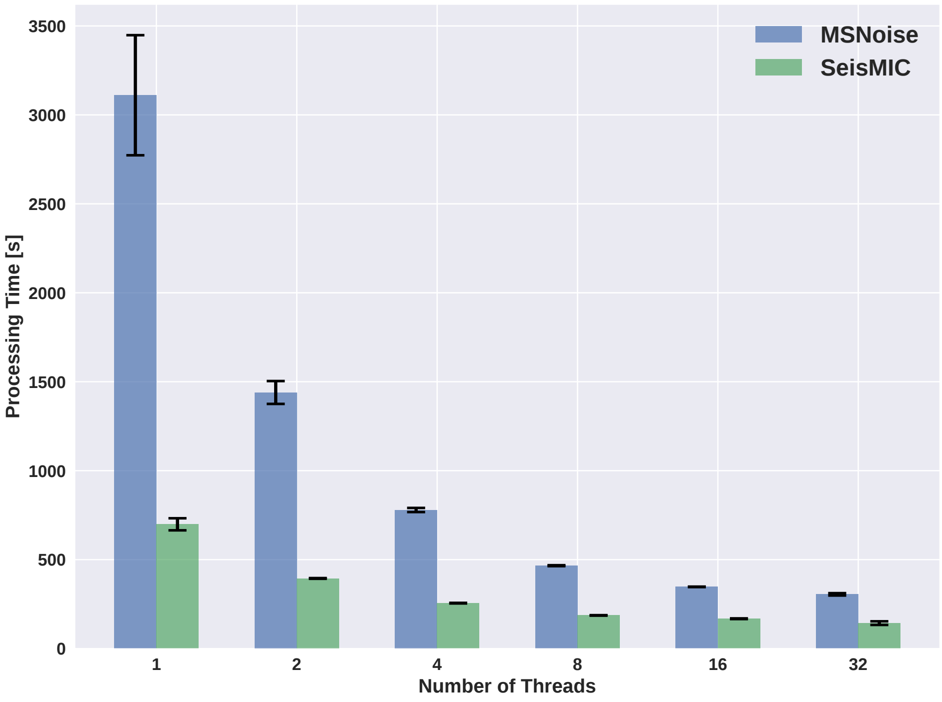


Figure 3 Compute times for a cross-correlation workflow for all six unique component combinations between eight seismic stations using MSNoise 1.6.3 (Lecocq et al., 2014) and SeisMIC 0.5.3. The height of the bars indicates the mean processing time over five iterations with the error bars representing the standard deviation. For hardware information and the exact parametrisation of the workflows, consult the text body.

2.3.2 Comparison with MSNoise

To analyse how SeisMIC's processing speed compares to the latest release of MSNoise (Lecocq et al., 2014), 1.6.3, we choose to calculate cross-correlations, which is the most expensive operation in a standard workflow, taking up more than 95% of the total compute time. In this benchmark, we retrieve hourly cross-correlations for 14 days of raw waveform data between eight 3-component broadband seismometers sampling at 100 Hz. We set the preprocessing to be identical for both programs. First, the data are decimated to 25 Hz. Subsequently, we detrend, taper, and band-pass filter the data between 2 and 4 Hz. Before computing the CFs, we apply one-bit normalisation and spectral whitening. We do not remove the instrument response. Note, however, that both MSNoise and SeisMIC execute the response removal using ObsPy (Beyreuther et al., 2010) and will therefore take the same amount of compute time and resources. Finally, we save the hourly CFs and daily CF stacks for all six unique component combinations with a length of 50 seconds. We perform the benchmark on the same Intel-Cascadelake-based node that we use in section 2.3.1.

We show the processing times required by MSNoise and SeisMIC for the outlined operation as a function of employed threads in Figure 3. Despite having received a significant performance boost with the update to version 1.6.x, MSNoise still needs about five times as long and thrice as much random access memory (RAM) as SeisMIC to execute the cross-correlation workflow, putting SeisMIC at a similar efficiency level as NoisePy (see Jiang and Denolle, 2020). In addition, SeisMIC offers a broader range of preprocessing options than NoisePy or MSNoise. MSNoise creates one miniseed file per CF, resulting in less complex writing operations, which are more evenly distributed across the cores. For this benchmark, this translates to a slightly better scaling between the number of cores and the computational time but also in a high number of files, which can be undesirable for large datasets. SeisMIC, on the other hand, creates one file per component combination. In every case, MSNoise remains more than twice as slow as SeisMIC. Note that the shown times do not include the time that MSNoise takes to set up a database and scan new data, which can take a significant amount of time, whereas these operations are practically instantaneous in SeisMIC.

While the presented results are encouraging, we remark that we could decrease compute times even further by exploiting the potential of modern graphic processing units (GPUs), which can correlate ambient seismic noise with high efficiency (Clements and Denolle, 2021; Wu et al., 2022). Implementing such algorithms belongs to the intermediate-term goals of SeisMIC's development.

3 A Practical Example of a Workflow: From Raw Waveform Data to a Velocity Change Time Series

In this section, we demonstrate how to obtain a dv/v time series using a minimal workflow in SeisMIC. In the supplementary material, we provide two Jupyter notebooks containing the source code used for this workflow. The exemplary data are recorded by station X9.IR1

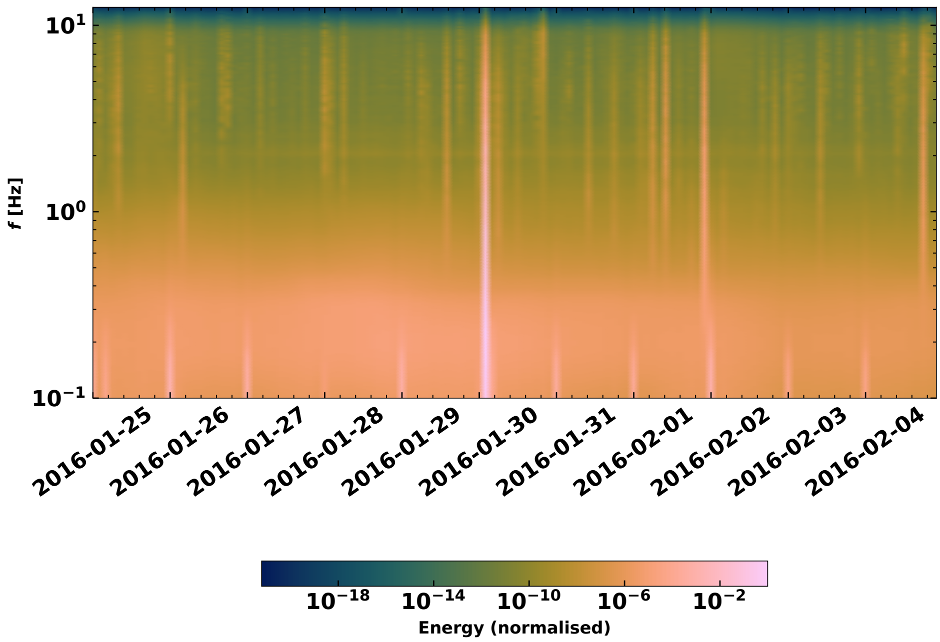


Figure 4 Time dependent spectrogram of the raw waveform at X9.IR1. We compute the spectrogram after removing the instrument response using 2 hours Welch windows. Note the energy spike caused by the Zhupanov earthquake. The energy amplitude is normalised by its maximum.

around the date of the M7.2 Zhupanov earthquake in Kamchatka, Russia. In the following, we investigate the impact of the event on the seismic velocity in the station's vicinity. A discussion of the result lies beyond the scope of this technical paper and has already been performed by Makus et al. (2023b). We conducted this analysis using SeisMIC's implemented workflow, which is parametrised using a simple *YAML* file (see supplementary material). In the following, we will take a step-by-step tour through said workflow and provide some minimal code examples. For further examples, we advise the reader to consult SeisMIC's documentation and our GitHub page.

3.1 Data Retrieval

To start, we download data from an FDSN-compatible server. In our case, we download data from station X9.IR1, available over the GEOFON FDSN service (Quinteros et al., 2021). For conciseness, we restrict this example to 11 days of data from 25 January to 5 February 2016. In section 2.3, we show how SeisMIC performs when confronted to larger datasets recorded on several stations and how compute time scales when employing multiple cores. Our exemplary time window comprises the 28 January Zhupanov earthquake, whose coseismic velocity drop we want to investigate. In SeisMIC, we can initiate the data download using the `Store_Client` class and its method `download_waveforms_mdl`:

Listing 1 Downloading data using SeisMIC

```
from obspy import UTCDateTime

from seismic.trace_data.waveform import
    Store_Client

starttime = UTCDateTime(2016, 1, 25)
endtime = UTCDateTime(2016, 2, 5)

# Decide where data are stored
sc = Store_Client('GEOFON', '/path/to/project
    ', read_only=False)
sc.download_waveforms_mdl(
    starttime, endtime, clients=['GEOFON'],
        network='X9',
    station='IR1', location='*', channel='HHE
        ')
```

Under the hood, this will initiate ObsPy's (Beyreuther et al., 2010) `MassDownloader` to download continuous waveform data from the specified station if not already present locally. Here, we will compute autocorrelations using only the east component of the seismogram. We can use SeisMIC to get a first idea of the spectral content of our waveform and to investigate in which frequency bands we might find stable noise sources suitable for PII. We show a spectrogram computed using Welch windows (see, e.g., Barbe et al., 2010) as implemented in SeisMIC in Figure 4.

3.2 Computing Autocorrelations

After downloading the waveforms, we can correlate them to obtain CFs. When computing correlations, we have ample preprocessing options, which, for brevity, we will not discuss here in detail. Most fundamentally, we must set the correlation length, `corr_len`, (i.e., the duration of the time windows to be correlated), the increment between these time windows, `corr_inc`, the correlation method (in our case, autocorrelation), and the frequency window to be filtered. The user defines all options in the *YAML* file, but they can also provide parameters in a Python dictionary. For this example, we choose a correlation length of one hour and a frequency band between 2 and 4 Hz. In SeisMIC, the `Correlator`

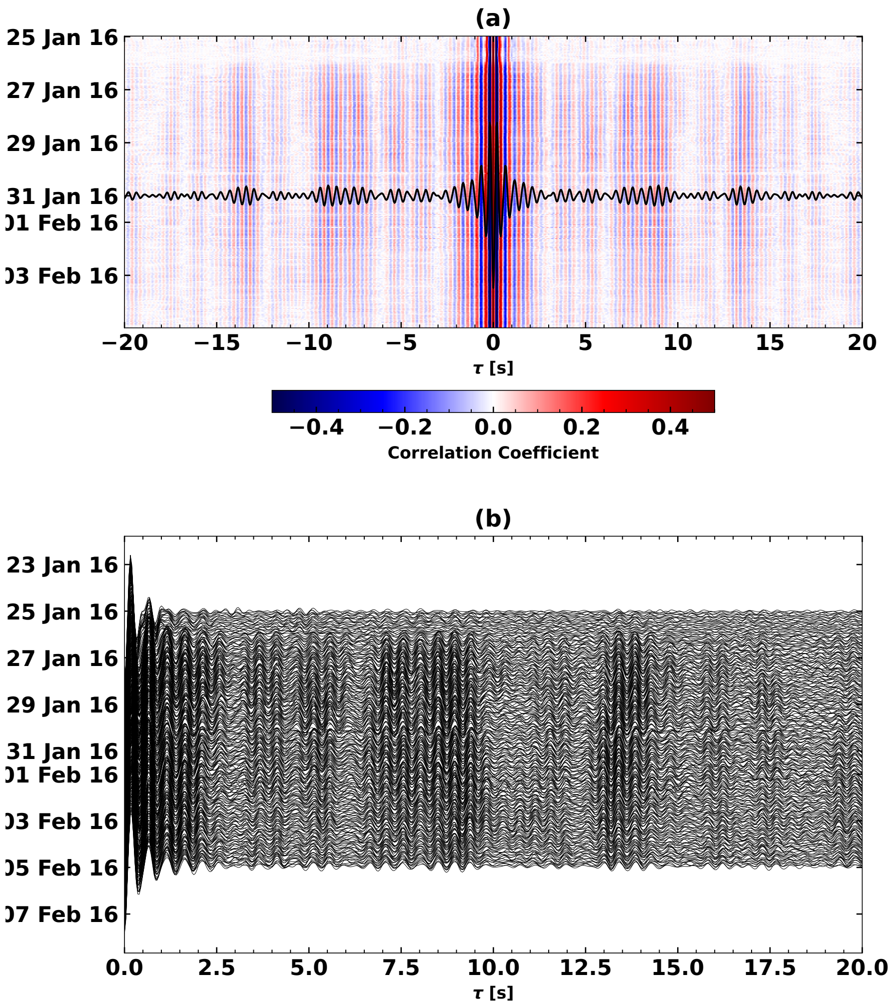


Figure 5 Hourly autocorrelations of ambient noise recorded by the east component of X9.IR1. This plot showcases two styles to plot correlations in SeisMIC. **(a)** Autocorrelations plotted as a colour image. The colours scale with the amplitude of the correlation. We superimpose the average of all shown autocorrelations on top of the heatmap. **(b)** Autocorrelations plotted as a section plot. In this plot, each hourly CF corresponds to one curve. Here, we only show the causal side of the CF.

class handles the correlation workflow.

Listing 2 Downloading data using SeisMIC

```
from seismic.correlate.correlate import
    Correlator

# sc is the previously initated Store_Client
c = Correlator(sc, options='path/to/params.
    YAML')
st = c.pxcorr()
```

To illustrate the syntax of the parameter file, we show an extract of it below. Note that the keys `preProcessing`, `TDpreProcessing`, and `FDpreProcessing` can also import custom, external functions as long as input arguments and return objects follow a predefined syntax.

Listing 3 params.YAML

```
...
read_start : '2016-01-25 00:00:01.0'
read_end : '2016-02-05 00:00:00.0'
sampling_rate : 25
remove_response : False
combination_method : 'autoComponents'
preProcessing : [
    {'function':'seismic.correlate.
        preprocessing_stream.detrend_st',
        'args':{'type':'linear'}},
    {'function':'seismic.correlate.
        preprocessing_stream.cos_taper_st',
        'args':{'taper_len': 100,
            'lossless': True}},
    {'function':'seismic.correlate.
        preprocessing_stream.stream_filter',
        'args':{'ftype':'bandpass',
            'filter_option':{'freqmin':0.01,
                'freqmax':12.49}}}]
subdivision:
    corr_inc : 3600
    corr_len : 3600
    ...
```

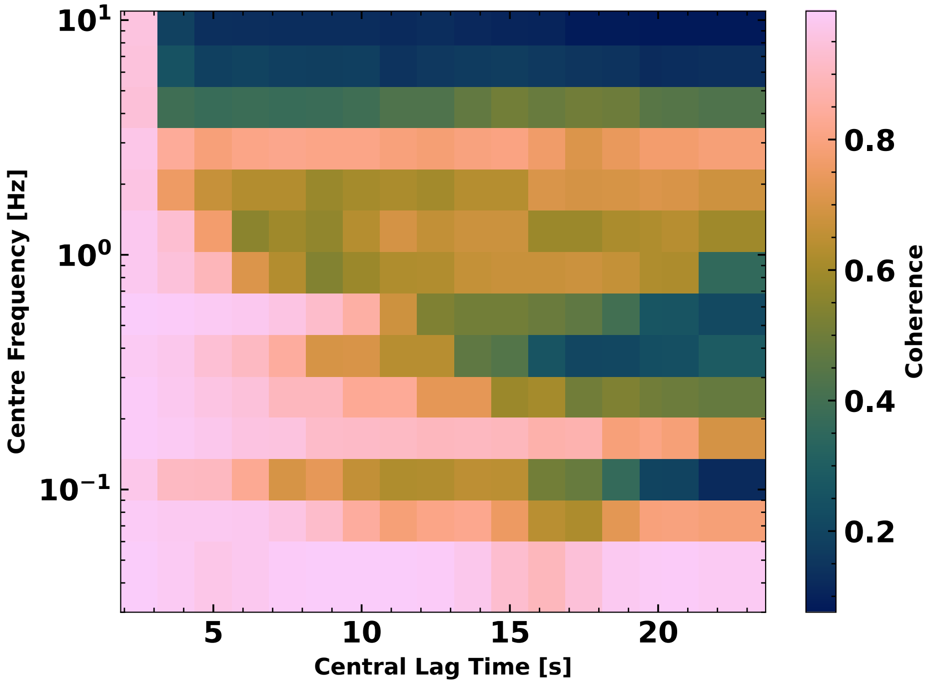


Figure 6 The waveform coherence as a function of lag time and frequency for the dataset from station X9.IR1 and channel HHE. For details, consult the text body.

```
corr_args : {'TDpreProcessing':[
    {'function':'seismic.correlate.
        preprocessing_td.detrend',
        'args':{'type':'linear'}},
    {'function':'seismic.correlate.
        preprocessing_td.TDfilter',
        'args':{'type':'bandpass','freqmin'
            :2,'freqmax':4}},
    ],
    'lengthToSave':25,
    'center_correlation':True,
    'normalize_correlation':True,
    ...
    }
        ...
```

Its `pxcorr` method will internally handle preprocessing and correlation. It will also initiate MPI to enable parallel processing. In Figure 5, we plotted the CFs using SeisMIC's plotting tools. Due to the high noise level in the chosen time window and frequency band, a well-defined coda emerges from the CFs (see Makus et al., 2023b, for details).

3.3 Waveform Coherence

For a first assessment of which frequency bands are well-suited for a velocity change analysis, we can use a spectrogram like the one we show in Figure 4. Additionally, one can use SeisMIC's waveform coherence function. The waveform coherence corresponds to the averaged zero-lag cross-correlation between a reference CF and CFs at time t (Steinmann et al., 2021). In Figure 6, we show the waveform coherence for our exemplary dataset computed between hourly CFs and the average CF as a reference. We determine the coherence for 5s long lapse-time windows and one-octave-wide frequency bands jointly for positive (causal) and negative (acausal) lag times. SeisMIC computes waveform coherence using the `Monitor` class and its `compute_waveform_coherence_bulk()` method (see supplementary material).

Figure 6 leads us to infer a high stability and energy content between 0.5 and 4 Hz. The coherence remains high until late lag times, e.g. for 3 Hz centre frequency, up to 75 periods. From this, we infer a highly scattering medium paired with a high energy content in this frequency band originating from the volcanic system (see Makus et al., 2023b). Therefore, we henceforth focus on the analysis of dv/v between 2 and 4 Hz.

3.4 Computing Velocity Changes Using the Stretching Method

Using the procedure theoretically outlined in section 2.2, we can estimate the evolution of the seismic velocity in the study period. Like previously, the parametrisation is handled over the *YAML* file (see supplementary material). Before computing dv/v, we smooth the one-hour CFs with a 4-hour long Hanning window. As reference CF, we use the mean of all CFs. Then, we compute dv/v for lag times between 3.5 s and 12 s simultaneously from the causal and acausal parts of the coda. We plot the resulting velocity change time series using one of SeisMIC's standard plotting templates in Figure 7.

Even though we do not focus on data interpretation in this article, we should take a brief look at the presented results. Most notably, we identify a clear velocity drop coinciding with the regional M7.2 Zhupanov earthquake. Interestingly, the resolution of the dv/v time series is high enough to identify a diurnal cycle that could be caused by air temperature and pressure variations, for example, observed by Wang et al. (2020), or might be due to lunar and solar tides as reported by Yamamura et al. (2003) and Sens-Schönfelder and Eulenfeld

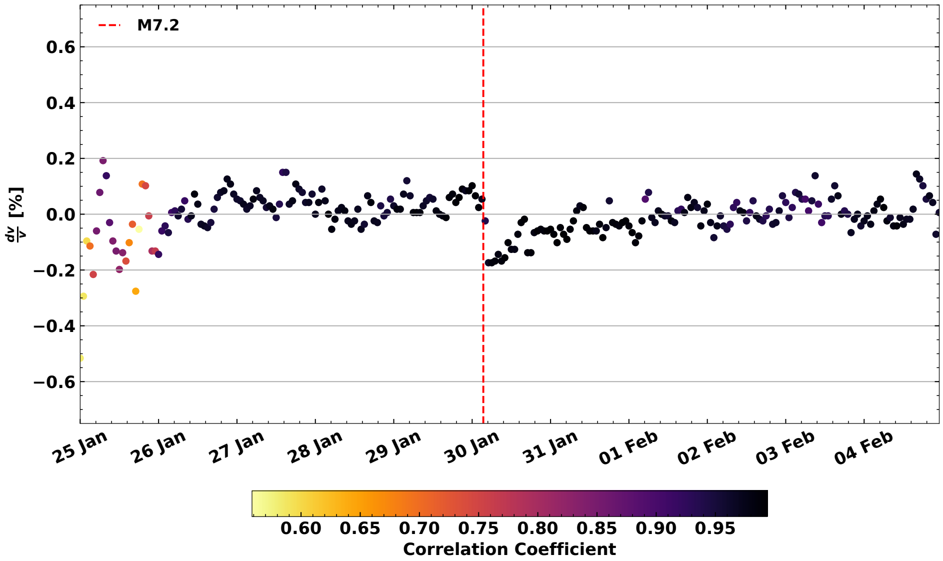


Figure 7 Velocity change time series estimated from the CFs shown in Figure 5. The increment between each data point is one hour and the shown dv/v is derived from CFs that are smoothed over 4 hours. The points' colour scales with the correlation coefficient (coherence) between the stretched CF and the reference CF. We plotted the origin time of the M7.2 Zhupanov earthquake, which occurred on 28 January 2016, as a vertical red line. An obvious velocity drop coinciding with the event can be identified. A subsequent recovery and more subtle differences in seismic velocity between day- and nighttime are visible.

(2019). Lastly, we note that the correlation coefficient is significantly lower before 26 January 2016. We link this observation to a transient change in the wavefield as described by Makus et al. (2023b) and Steinmann et al. (2023).

3.5 Spatial Imaging of Velocity Changes

Velocity change estimates like the one presented in Figure 7 show dv/v as a function of time but do not directly yield insight into the spatial distribution of these velocity changes. Coda waves, as used in PII, sample the medium at a high spatial extent. While this allows to detect distributed weak velocity changes or changes located away from the path of direct waves, it prevents a simple inference of the affected location along a ray path or Fresnel volume. The affected location can, however, be estimated using sensitivity kernels that describe the time-dependent energy distribution of the wavefield for a statistically uniform medium. For a theoretical derivation of the sensitivity kernels based on the Radiative Transfer Theory, refer to Mayor et al. (2014), Margerin et al. (2016), and Zhang et al. (2022).

In SeisMIC, we implemented a simplified approach relying on sensitivity kernels derived from an approximate solution of the Boltzmann equation for a homogeneous medium (Paasschens, 1997) describing isotropic scattering of acoustic waves. Using these sensitivity kernels and a linearised inversion scheme proposed by Obermann et al. (2013), we can map a 2-dimensional distribution of dv/v at a fixed time t_i resulting in $dv/v(t_i, x, y)$.

In SeisMIC, the module `seismic.monitor.spatial` contains the necessary functions for the outlined approach. To illustrate the procedure and make our example easily adaptable and reproducible, we create a synthetic velocity-change model, which we then forward model onto a random station configuration. After adding noise to the synthetic data, we try to recover the initial model using the inverse algorithm. In detail, we proceed as follows: First, we create a synthetic velocity change model with an extent of 40 km×40 km and a spatial resolution of 1 km (Figures 8 (b) and (d)). The background medium has a homogeneous velocity of 3 $\frac{\mathrm{km}}{\mathrm{s}}$ and a transport mean free path l_0 of 30 km. Then, we place an arbitrary number of stations on random positions along the grid. Using sensitivity kernels of cross- and autocorrelations, we solve the forward problem to compute dv/v, as it would be obtained from the CFs in the presence of the spatial velocity variations. The sensitivity kernels are computed for lapse time windows between 14 and 34 s. To the dv/v values, we add random noise. This noise follows a Gaussian distribution around 0% velocity change with a standard deviation of 0.1%. Finally, we invert for the synthetic model employing the damped linearised inversion (Tarantola and Valette, 1982). We show the results of this inversion in Figures 8 (a) and (c) for 4 and 32 stations, respectively. There, we also indicate the used damping parameters. The optimal damping parameters minimise both the misfit between the initial and the retrieved model and the model complexity and can be found using the L-curve criterion, as discussed by Obermann et al. (2013). This inversion relies on two damping parameters, the correlation length λ determining how strongly related neighbouring grid cells are and the model variance σ_m that the model may assume.

The results demonstrate that increasing the number of stations is the most powerful tool to decrease the misfit between the inversion result and the input model.

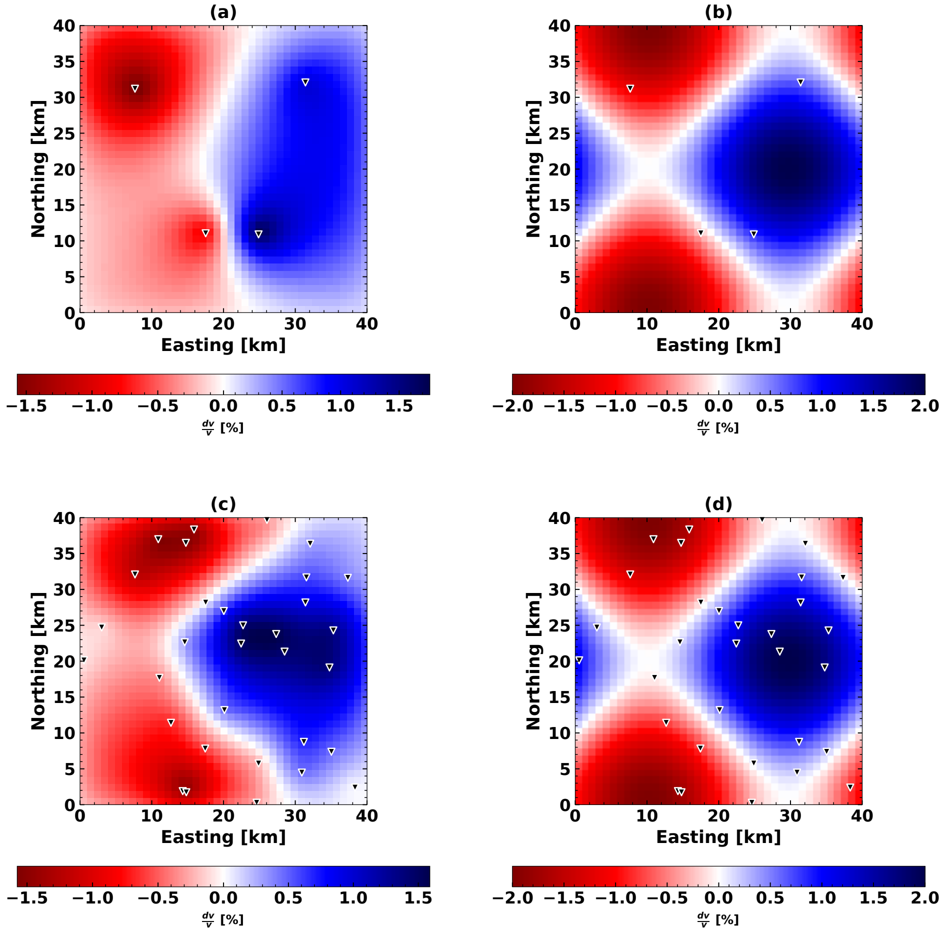


Figure 8 Two examples of the spatial inversion using different parametrisations and station configurations.**(a)** Result of the spatial inversion algorithm using four stations, a model variance $\sigma_m = 0.1\,\frac{\mathrm{km}}{\mathrm{km}^2}$, and a correlation length $\lambda = 2$ km. **(b)** The synthetic velocity model and station configuration used to obtain (a). **(c)** Result of the spatial inversion algorithm using 32 stations, $\sigma_m = 0.01\,\frac{\mathrm{km}}{\mathrm{km}^2}$, and $\lambda = 2$ km. **(d)** The synthetic velocity model and station configuration used to obtain (c). For an exhaustive description of the parametrisation and the inversion steps, consult the text body.

While the geometry of the synthetic model is poorly retrieved for a configuration using only four stations, we can reproduce the model quite accurately with 32 stations.

The supplementary material contains a Jupyter notebook to reproduce or modify these results with an arbitrary number of stations, velocity change model, and damping parameters. We also include options to invert for dv/v only utilising data from auto- or cross-correlations and using sensitivity kernels from split coda windows (i.e., with lapse time windows sliced into narrow sub-windows). In the supplement, we show results that exploit these options. Based on these, we argue that adding dv/v information from auto- and cross-correlations, improves the accuracy of the result notably, whereas splitting the coda yields only minor improvements.

4 Conclusion and Outlook

We presented SeisMIC, a software to estimate changes in the seismic propagation velocity from continuous records of seismic ambient noise. SeisMIC contains functionalities for the end-to-end processing of velocity-change time series, including data retrieval, the computation of correlation functions, calculating velocity change time series using the stretching method, and postprocessing as well as inverting dv/v time series onto a spatial grid. While these functions can be part of a workflow, they are also intended to be used separately and can easily be altered and adapted to individual processes. In SeisMIC, we implement a new data format for correlation functions, which provides uniformity, flexibility, interoperability, and integrity. Thereby, we hope to foster a broader discussion in the community regarding data standards, which, we believe, would aid data exchange, efficiency, and reproducibility of ambient noise studies.

In the near future, we will release versions capable of estimating dv/v employing algorithms other than the stretching method, like the wavelet-cross-spectrum analysis (Mao et al., 2020). Other future milestones include exploiting the computational power of GPUs to decrease the compute time of noise correlations even further and adding solutions that automatically update correlation function databases.

SeisMIC complements existing software to process ambient noise. Highlights are its broad functionality, high efficiency, and versatility applicable to local small-scale studies on a laptop computer as well as surveys using large-N arrays processed on computer clusters. SeisMIC is available on GitHub as a well-documented and regularly maintained open-source software.

Acknowledgements

We thank Luc Illien, Ugo Nanni, Jannes Münchmeyer, Marine Denolle, Charlotte Bruland, Tuo Zhang, and Raphael De Plaen for discussions that helped improving this software. This work was supported by the DFG grant SE1844/12-1. We are grateful for comments and suggestions towards improvement of this work by Laura Ermert and two anonymous reviewers.

Data and code availability

This manuscript is distributed with a supporting information in PDF format. Aside from the supporting document, we provide Jupyter notebooks, computing scripts, and the main program "SeisMIC" as a digital supplement. For SeisMIC, however, we strongly encourage the reader to obtain the code's latest version, for example, from GitHub. The digital supplement includes SeisMIC 0.5.3 and is available at https://doi.org/10.5281/zenodo.8283683.

The SeisMIC data from the KISS experiment (Shapiro et al., 2017) used in section 3 can be obtained from the GEOFON webservice (Quinteros et al., 2021). For the benchmarks in section 2.3, we used data from the Saxony Network (University of Leipzig, 2001) available from the EIDA FDSN service (Strollo et al., 2021).

Competing interests

The authors do not declare any competing interests.

References

Arrowsmith, S. J., Trugman, D. T., MacCarthy, J., Bergen, K. J., Lumley, D., and Magnani, M. B. Big Data Seismology. *Reviews of Geophysics*, 60(2):e2021RG000769, 2022. doi: 10.1029/2021RG000769.

Asnar, M., Sens-Schönfelder, C., Bonnelye, A., and Dresen, G. Non-Classical, Non-Linear Elasticity in Rocks: Experiments in a Triaxial Cell with Pore Pressure Control. Technical Report EGU23-8050, Copernicus Meetings, Feb. 2023. doi: 10.5194/egusphere-egu23-8050.

Barbe, K., Pintelon, R., and Schoukens, J. Welch Method Revisited: Nonparametric Power Spectrum Estimation Via Circular Overlap. *IEEE Transactions on Signal Processing*, 58(2):553–565, Feb. 2010. doi: 10.1109/TSP.2009.2031724.

Barker, M., Chue Hong, N. P., Katz, D. S., Lamprecht, A.-L., Martinez-Ortiz, C., Psomopoulos, F., Harrow, J., Castro, L. J., Gruenpeter, M., Martinez, P. A., and Honeyman, T. Introducing the FAIR Principles for Research Software. *Scientific Data*, 9(1):622, Oct. 2022. doi: 10.1038/s41597-022-01710-x.

Bensen, G. D., Ritzwoller, M. H., Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., Shapiro, N. M., and Yang, Y. Processing Seismic Ambient Noise Data to Obtain Reliable Broad-Band Surface Wave Dispersion Measurements. *Geophysical Journal International*, 169(3):1239–1260, 2007. doi: 10.1111/j.1365-246X.2007.03374.x.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., and Wassermann, J. ObsPy: A Python Toolbox for Seismology. *Seismological Research Letters*, 81(3):530–533, May 2010. doi: 10.1785/gssrl.81.3.530.

Bièvre, G., Franz, M., Larose, E., Carrière, S., Jongmans, D., and Jaboyedoff, M. Influence of Environmental Parameters on the Seismic Velocity Changes in a Clayey Mudflow (Pont-Bourquin Landslide, Switzerland). *Engineering Geology*, 245:248–257, Nov. 2018. doi: 10.1016/j.enggeo.2018.08.013.

Brenguier, F., Campillo, M., Hadziioannou, C., Shapiro, N. M., Nadeau, R. M., and Larose, E. Postseismic Relaxation Along the San Andreas Fault at Parkfield from Continuous Seismological Observations. *Science*, 321(5895):1478–1481, Sept. 2008. doi: 10.1126/science.1160943.

Clements, T. and Denolle, M. A. Tracking Groundwater Levels Using the Ambient Seismic Field. *Geophysical Research Letters*, 45(13): 6459–6465, 2018. doi: 10.1029/2018GL077706.

Clements, T. and Denolle, M. A. SeisNoise.Jl: Ambient Seismic Noise Cross Correlation on the CPU and GPU in Julia. *Seismological Research Letters*, 92(1):517–527, 2021. doi: 10.1785/0220200192.

Clymer, R. W. and McEvilly, T. V. Travel-Time Monitoring with Vibroseis*. *Bulletin of the Seismological Society of America*, 71(6): 1903–1927, Dec. 1981. doi: 10.1785/BSSA0710061903.

Collette, A., Kluyver, T., Caswell, T. A., Tocknell, J., Kieffer, J., Scopatz, A., Dale, D., Chen, Jelenak, A., payno, juliagarriga, VINCENT, T., Sciarelli, P., Valls, V., Kofoed Pedersen, U., jakirkham, Raspaud, M., Parsons, A., Abbasi, H., Readey, J., Paramonov, A., Chan, L., jialin, Danilevski, C., Feng, Y., Vaillant, G. A., Teichmann, M., Brucher, M., Johnson, S. R., and Hole, L. H5py/H5py: 3.1.0. *Zenodo*, Nov. 2020. doi: 10.5281/zenodo.4250762.

Dalcin, L. and Fang, Y.-L. L. Mpi4py: Status Update After 12 Years of Development. *Computing in Science & Engineering*, 23(4):47–54, July 2021. doi: 10.1109/MCSE.2021.3083216.

Donaldson, C., Winder, T., Caudron, C., and White, R. S. Crustal Seismic Velocity Responds to a Magmatic Intrusion and Seasonal Loading in Iceland's Northern Volcanic Zone. *Science Advances*, 5(11):eaax6642, Nov. 2019. doi: 10.1126/sciadv.aax6642.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E. Array Programming with NumPy. *Nature*, 585(7825):357–362, Sept. 2020. doi: 10.1038/s41586-020-2649-2.

Hirose, T., Nakahara, H., and Nishimura, T. Combined Use of Repeated Active Shots and Ambient Noise to Detect Temporal Changes in Seismic Velocity: Application to Sakurajima Volcano, Japan 4. Seismology. *Earth, Planets and Space*, 69(1), 2017. doi: 10.1186/s40623-017-0613-7.

Hong, N. P. C., Katz, D. S., Barker, M., Lamprecht, A.-L., Martinez, C., Psomopoulos, F. E., Harrow, J., Castro, L. J., Gruenpeter, M., Martinez, P. A., Honeyman, T., Struck, A., Lee, A., Loewe, A., van Werkhoven, B., Garijo, D., Plomp, E., Genova, F., Shanahan, H., Hellström, M., Sandström, M., Sinha, M., Kuzak, M., Herterich, P., Islam, S., Sansone, S.-A., Pollard, T., Atmojo, U. D., Williams, A., Czerniak, A., Niehues, A., Fouilloux, A. C., Desinghu, B., Goble, C., Richard, C., Gray, C., Erdmann, C., Nüst, D., Tartarini, D., Ranguelova, E., Anzt, H., Todorov, I., McNally, J., Burnett, J., Garrido-Sánchez, J., Belhajjame, K., Sesink, L., Hwang, L., Tovani-Palone, M. R., Wilkinson, M. D., Servillat, M., Liffers, M., Fox, M., Miljković, N., Lynch, N., Lavanchy, P. M., Gesing, S., Stevens, S., Cuesta, S. M., Peroni, S., Soiland-Reyes, S., Bakker, T., Rabemanantsoa, T., Sochat, V., Yehudi, Y., and Wg, F. FAIR Principles for Research Software (FAIR4RS Principles). Mar. 2022. doi: 10.15497/RDA00065.

Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, 9(03):90–95, May 2007. doi: 10.1109/MCSE.2007.55.

Ikuta, R., Yamaoka, K., Miyakawa, K., Kunitomo, T., and Kumazawa, M. Continuous Monitoring of Propagation Velocity of Seismic Wave Using ACROSS. *Geophysical Research Letters*, 29(13): 5–1–5–5, 2002. doi: 10.1029/2001GL013974.

Illien, L., Andermann, C., Sens-Schönfelder, C., Cook, K. L., Baidya, K. P., Adhikari, L. B., and Hovius, N. Subsurface Moisture Regulates Himalayan Groundwater Storage and Discharge. *AGU Advances*, 2(2), 2021. doi: 10.1029/2021av000398.

Jiang, C. and Denolle, M. A. Noisepy: A New High-Performance Python Tool for Ambient-Noise Seismology. *Seismological Research Letters*, 91(3):1853–1866, 2020. doi: 10.1785/0220190364.

Koranne, S. Hierarchical Data Format 5 : HDF5. In Koranne, S., editor, *Handbook of Open Source Tools*, pages 191–200. Springer US, Boston, MA, 2011.

Lecocq, T., Caudron, C., and Brenguier, F. MSNoise, a Python Package for Monitoring Seismic Velocity Changes Using Ambient Seismic Noise. *Seismological Research Letters*, 85(3):715–726, May 2014. doi: 10.1785/0220130073.

Lindner, F., Wassermann, J., and Igel, H. Seasonal Freeze-Thaw Cycles and Permafrost Degradation on Mt. Zugspitze (German/Austrian Alps) Revealed by Single-Station Seismic Monitoring. *Geophysical Research Letters*, 48(18):1–11, 2021. doi: 10.1029/2021GL094659.

Makus, P. and Sens-Schönfelder, C. Seismological Monitoring Using Interferometric Concepts (SeisMIC). GFZ Data Services, 2022. doi: 10.5880/GFZ.2.4.2022.002.

Makus, P., Denolle, M., Sens-Schönfelder, C., Köpfli, M., and Tilmann, F. The Complex Relationship between Seismic Velocity and Volcanic, Tectonic, and Environmental Forcings Illustrated by 23 Years of Data at Mt. St. Helens. In *EGU23*. Copernicus Meetings, Feb. 2023a. doi: 10.5194/egusphere-egu23-6757.

Makus, P., Sens-Schönfelder, C., Illien, L., Walter, T. R., Yates, A., and Tilmann, F. Deciphering the Whisper of Volcanoes: Monitoring Velocity Changes at Kamchatka's Klyuchevskoy Group With Fluctuating Noise Fields. *Journal of Geophysical Research: Solid Earth*, 128(4):e2022JB025738, 2023b. doi: 10.1029/2022JB025738.

Mao, S., Mordret, A., Campillo, M., Fang, H., and van der Hilst, R. D. On the Measurement of Seismic Traveltime Changes in the Time–Frequency Domain with Wavelet Cross-Spectrum Analysis. *Geophysical Journal International*, 221(1):550–568, Apr. 2020. doi: 10.1093/gji/ggz495.

Mao, S., Lecointre, A., van der Hilst, R. D., and Campillo, M. Space-Time Monitoring of Groundwater Fluctuations with Passive Seismic Interferometry. *Nature Communications*, 13(1): 4643, Aug. 2022. doi: 10.1038/s41467-022-32194-3.

Margerin, L., Planès, T., Mayor, J., and Calvet, M. Sensitivity Kernels for Coda-Wave Interferometry and Scattering Tomography: Theory and Numerical Evaluation in Two-Dimensional Anisotropically Scattering Media. *Geophysical Journal International*, 204(1):650–666, Jan. 2016. doi: 10.1093/gji/ggv470.

Mayor, J., Margerin, L., and Calvet, M. Sensitivity of Coda Waves to Spatial Variations of Absorption and Scattering: Radiative Transfer Theory and 2-D Examples. *Geophysical Journal International*, 197(2):1117–1137, May 2014. doi: 10.1093/gji/ggu046.

Minato, S., Tsuji, T., Ohmi, S., and Matsuoka, T. Monitoring Seismic Velocity Change Caused by the 2011 Tohoku-oki Earthquake Using Ambient Noise Records. *Geophysical Research Letters*, 39(9), 2012. doi: 10.1029/2012GL051405.

Mordret, A., Mikesell, T. D., Harig, C., Lipovsky, B. P., and Prieto, G. A. Monitoring Southwest Greenland's Ice Sheet Melt with Ambient Seismic Noise. *Science Advances*, 2(5):e1501538, May 2016. doi: 10.1126/sciadv.1501538.

Nanni, U., Pauze, T., Goulet, L., Köhler, A., Bouchayer, C., and Schuler, T. Study of the Structural and Dynamic Changes of a Surging Glacier Using Seismic Observations. In *IUGG*, Berlin, July 2023.

Nishimura, T., Uchida, N., Sato, H., Ohtake, M., Tanaka, S., and Hamaguchi, H. Temporal Changes of the Crustal Structure Associated with the M6.1 Earthquake on September 3, 1998, and the Volcanic Activity of Mount Iwate, Japan. *Geophysical Research Letters*, 27(2):269–272, 2000. doi: 10.1029/1999GL005439.

Obermann, A., Planès, T., Larose, E., and Campillo, M. Imaging Preeruptive and Coeruptive Structural and Mechanical Changes of a Volcano with Ambient Seismic Noise. *Journal of Geophysical Research: Solid Earth*, 118(12):6285–6294, 2013. doi: 10.1002/2013JB010399.

Paasschens, J. C. J. Solution of the Time-Dependent Boltzmann Equation. *Physical Review E*, 56(1):1135–1141, July 1997. doi: 10.1103/PhysRevE.56.1135.

Poupinet, G., Ellsworth, W. L., and Frechet, J. Monitoring Velocity Variations in the Crust Using Earthquake Doublets: An Application to the Calaveras Fault, California. *Journal of Geophysical Research: Solid Earth*, 89(B7):5719–5731, July 1984. doi: 10.1029/JB089iB07p05719.

Quinteros, J., Strollo, A., Evans, P. L., Hanka, W., Heinloo, A., Hemmleb, S., Hillmann, L., Jaeckel, K.-H., Kind, R., Saul, J., Zieke, T., and Tilmann, F. The GEOFON Program in 2020. *Seismological Research Letters*, 92(3):1610–1622, Feb. 2021. doi: 10.1785/0220200415.

Sawazaki, K., Kimura, H., Shiomi, K., Uchida, N., Takagi, R., and Snieder, R. Depth-Dependence of Seismic Velocity Change Associated with the 2011 Tohoku Earthquake, Japan, Revealed from Repeating Earthquake Analysis and Finite-Difference Wave Propagation Simulation. *Geophysical Journal International*, 201 (2):741–763, May 2015. doi: 10.1093/gji/ggv014.

Sens-Schönfelder, C. and Eulenfeld, T. Probing the in Situ Elastic Nonlinearity of Rocks with Earth Tides and Seismic Noise. *Physical Review Letters*, 122(13):138501, Apr. 2019. doi: 10.1103/PhysRevLett.122.138501.

Sens-Schönfelder, C. and Wegler, U. Passive Image Interferemetry and Seasonal Variations of Seismic Velocities at Merapi Volcano, Indonesia. *Geophysical Research Letters*, 33(21):1–5, 2006. doi: 10.1029/2006GL027797.

Sens-Schönfelder, C. and Wegler, U. Passive Image Interferometry for Monitoring Crustal Changes with Ambient Seismic Noise. *Comptes Rendus Geoscience*, 343(8):639–651, Sept. 2011. doi: 10.1016/j.crte.2011.02.005.

Sens-Schönfelder, C., Flores-Estrella, H., Gassenmeier, M., Korn, M., Köllner, F., Milkereit, C., Niederleithinger, E., Parolai, S., Pilz, M., Pomponi, E., Schuck, A., Thiemann, K., and Völkel, J. MIIC: Monitoring and Imaging Based on Interferometric Concepts. In Weber, M. and Münch, U., editors, *Tomography of the Earth's Crust: From Geophysical Sounding to Real-Time Monitoring: GEOTECHNOLOGIEN Science Report No. 21*, Advanced Technologies in Earth Sciences, pages 43–61. Springer International Publishing, Cham, 2014a.

Sens-Schönfelder, C., Pomponi, E., and Peltier, A. Dynamics of Piton de La Fournaise Volcano Observed by Passive Image Interferometry with Multiple References. *Journal of Volcanology and Geothermal Research*, 276:32–45, 2014b. doi: 10.1016/j.jvolgeores.2014.02.012.

Shapiro, N., Sens-Schönfelder, C., Lühr, B., Weber, M., Abkadyrov, I., Gordeev, E., Koulakov, I., Jakovlev, A., Kugaenko, Y., and Saltykov, V. Understanding Kamchatka's Extraordinary Volcano Cluster. *Eos*, May 2017. doi: 10.1029/2017EO071351.

Snieder, R., Grêt, A., Douma, H., and Scales, J. Coda Wave Interferometry for Estimating Nonlinear Behavior in Seismic Velocity. *Science*, 295(5563):2253–2255, 2002. doi: 10.1126/science.1070015.

Steinmann, R., Hadziioannou, C., and Larose, E. Effect of Centimetric Freezing of the near Subsurface on Rayleigh and Love Wave Velocity in Ambient Seismic Noise Correlations. *Geophysical Journal International*, 224(1):626–636, 2021. doi: 10.1093/gji/ggaa406.

Steinmann, R., Seydoux, L., Journeau, C., Shapiro, N. M., and Campillo, M. Machine Learning Analysis of Seismograms Reveals a Continuous Plumbing System Evolution beneath the Klyuchevskoy Volcano in Kamchatka, Russia, June 2023. doi: 10.22541/essoar.168614505.54607219/v1.

Strollo, A., Cambaz, D., Clinton, J., Danecek, P., Evangelidis, C. P., Marmureanu, A., Ottemöller, L., Pedersen, H., Sleeman, R., Stammler, K., Armbruster, D., Bienkowski, J., Boukouras, K., Evans, P. L., Fares, M., Neagoe, C., Heimers, S., Heinloo, A., Hoffmann, M., Kaestli, P., Lauciani, V., Michalek, J., Odon Muhire, E., Ozer, M., Palangeanu, L., Pardo, C., Quinteros, J., Quintiliani, M., Antonio Jara-Salvador, J., Schaeffer, J., Schloemer, A., and Triantafyllis, N. EIDA: The European Integrated Data Archive and Service Infrastructure within ORFEUS. *Seismological Research Letters*, 92(3):1788–1795, Mar. 2021. doi: 10.1785/0220200413.

Tarantola, A. and Valette, B. Generalized Nonlinear Inverse Problems Solved Using the Least Squares Criterion. *Reviews of Geophysics*, 20(2):219–232, 1982. doi: 10.1029/RG020i002p00219.

University of Leipzig. SXNET Saxon Seismic Network, 2001. doi: 10.7914/SN/SX.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., and van Mulbregt, P. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17(3):261–272, Mar. 2020. doi: 10.1038/s41592-019-0686-2.

Wang, B., Yang, W., Wang, W., Yang, J., Li, X., and Ye, B. Diurnal and Semidiurnal P- and S-Wave Velocity Changes Measured Using an Airgun Source. *Journal of Geophysical Research: Solid Earth*, 125 (1):e2019JB018218, 2020. doi: 10.1029/2019JB018218.

Wang, Q. Y., Brenguier, F., Campillo, M., Lecointre, A., Takeda, T., and Aoki, Y. Seasonal Crustal Seismic Velocity Changes Throughout Japan. *Journal of Geophysical Research: Solid Earth*, 122(10):7987–8002, 2017. doi: 10.1002/2017JB014307.

Wegler, U., Lühr, B. G., Snieder, R., and Ratdomopurbo, A. Increase of Shear Wave Velocity before the 1998 Eruption of Merapi Volcano (Indonesia). *Geophysical Research Letters*, 33(9):4–7, 2006. doi: 10.1029/2006GL025928.

Wu, C., Tan, X., Li, H., and Sun, G. An Efficient Ambient Noise Cross-Correlation Algorithm on Heterogeneous CPU-GPU Cluster. In *2022 IEEE 13th International Symposium on Parallel Architectures, Algorithms and Programming (PAAP)*, pages 1–5, Nov. 2022. doi: 10.1109/PAAP56126.2022.10010612.

Yamamura, K., Sano, O., Utada, H., Takei, Y., Nakao, S., and Fukao, Y. Long-Term Observation of in Situ Seismic Velocity and Attenuation. *Journal of Geophysical Research: Solid Earth*, 108(B6), 2003. doi: 10.1029/2002JB002005.

Yang, W., Wang, B., Yuan, S., and Ge, H. Temporal Variation of Seismic-Wave Velocity Associated with Groundwater Level Observed by a Downhole Airgun near the Xiaojiang Fault Zone. *Seismological Research Letters*, 89(3):1014–1022, Apr. 2018. doi: 10.1785/0220170282.

Zhang, H., Glasgow, M., Schmandt, B., Thelen, W. A., Moran, S. C., and Thomas, A. M. Revisiting the Depth Distribution of Seismicity before and after the 2004–2008 Eruption of Mount St. Helens. *Journal of Volcanology and Geothermal Research*, 430:107629, Oct. 2022. doi: 10.1016/j.jvolgeores.2022.107629.

REPORT
doi:10.26443/seismica.v3i1.1049

SEISMICA

Curated Regional Earthquake Waveforms (CREW) Dataset

Albert L. Aguilar Suarez * [1], **Gregory C. Beroza** [1]

[1]Department of Geophysics, Stanford University, Stanford, CA, USA

Author contributions: *Conceptualization*: A.L. Aguilar S., G.C. Beroza. *Software*: A.L. Aguilar S.. *Writing - original draft*: A.L. Aguilar S., G.C. Beroza. *Writing - Review & Editing*: A.L. Aguilar S., G.C. Beroza. *Funding acquisition*: G.C. Beroza.

Abstract We have assembled CREW, the Curated Regional Earthquake Waveforms Dataset, which is a dataset of earthquake arrivals recorded at local and regional distances. CREW was assembled from millions of waveforms with quality control through semi-supervised learning. CREW includes 1.6 million waveforms that have global coverage. Each waveform consists of a 5 minute three component seismogram with labels for both a P and S arrival. CREW provides a high quality labeled waveform data set that can be used to develop and test machine learning models for the analysis of earthquakes recorded at regional distances.

Production Editor:
Carmine Galasso
Handling Editor:
Jack Muir
Copy & Layout Editor:
Théa Ragon

Received:
June 21, 2023
Accepted:
April 20, 2024
Published:
May 24, 2024

1 Introduction

The Deep Learning seismological data landscape is dominated by local recordings. STEAD (Mousavi et al., 2019) contains over 1.2 million three component earthquake waveforms recorded at distances up to 350 km, with 8 percent of the data recorded at more than 110 km. STEAD provides 60 s waveforms from around the world that include both P and S arrival labels. INSTANCE (Michelini et al., 2021) contains over 1.1 million three component earthquake waveforms recorded at distances up to 600 km. INSTANCE provides 120 s waveforms from Italy and its surroundings with at least a P or S arrival. LENDB (Magrini et al., 2020) contains over 600,000 three component earthquake waveforms recorded at distances up to 134 km. LENDB provides 27 s waveforms from around the world with picked P arrivals. The Pacific Northwest AI-ready Seismic Dataset (Ni et al., 2023) contains 190,000 three component waveforms for earthquakes and exotic events. This dataset provides 150 s waveforms. These four datasets also contain noise waveforms. The NEIC dataset (Yeck et al., 2020) contains over 1.3 million earthquake waveforms recorded at distances up to 90 degrees. This dataset provides 60 seconds long waveforms around the phases P, Pn, Pg, Sn, Sg and S, with the majority corresponding to P phases. The MLAAPDE dataset (Cole et al., 2023) contains 5.1 million three component waveforms for earthquakes recorded at distances ranging from local to teleseismic. This dataset provides 120 s waveforms. The GEOFON dataset (Woollam et al., 2022) also covers the local to teleseismic distance range, with nearly 275K labeled arrivals, mostly P waves.

Most seismological deep learning research on earthquake detection and phase picking has used short duration waveforms from small earthquakes at short distances. PhaseNet (Zhu and Beroza, 2019) was trained on 30 second waveforms to predict the timing of P and S wave arrivals in Northern California. Earthquake Transformer (Mousavi et al., 2020) was trained on 60 s waveforms to simultaneously detect earthquakes and pick the arrival times of P and S waves. (Woollam et al., 2019) used 6 s windows for phase picking and (Ross et al., 2018) employed 4 s windows to predict the type of dominant energy in the seismogram (P or S), training on seismograms recorded within 100 km from the epicenter.

Most of the world is sparsely instrumented and many earthquakes are recorded only at distances over 100 km. This is true for the important case of seismicity near most subduction trenches, which are often more than 100 km from the nearest land. At regional distances, which are often taken to be more than 100 km and up to 1,000 km, seismic waves are strongly modified by interaction of the wavefield with the crust and upper mantle. S and P arrivals are also separated by greater times than for shorter distances, such that existing deep learning models may not perform well on these out of distribution data. This provides the motivation for developing CREW. The increase in source-receiver distance comes with mounting complexity in the waveforms due to the accumulation of propagation effects and the decrease in wave amplitudes. Figure 1 schematically compares wave propagation at local distances vs. regional distances.

The waveforms on top are recordings of the 2023 Lake Almanor earthquake in Northern California. This M_W

*Corresponding author: aguilars@stanford.edu

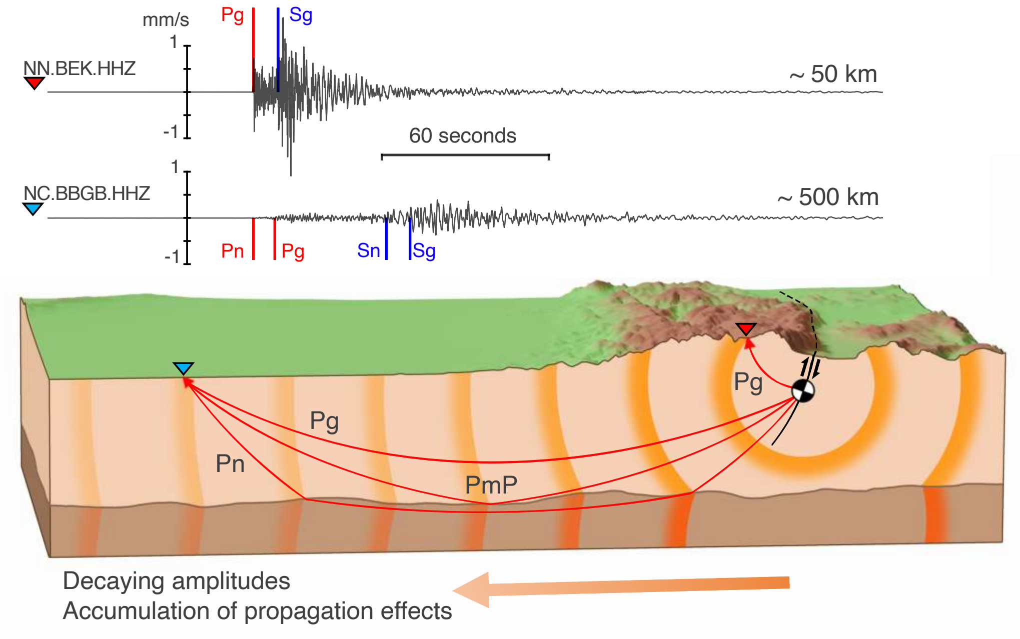


Figure 1 Comparison of local and regional recordings for the 2023 Lake Almanor earthquake in Northern California. The top waveform was recorded at around 50 km while the bottom one was recorded at about 500 km. Note: The focal mechanism, elevation and fault geometry are not related to the real Lake Almanor earthquake setting.

5.5 earthquake was recorded over many instruments at both distance ranges. The top seismogram comes from station BEK from the Nevada Seismological Network at a distance of around 50 km. In this case the arrivals of the crustal phases Pg and Sg are very impulsive and they are around 10 seconds apart. The bottom waveform, from station BBGB of the Northern California Seismic Network recorded the earthquake at a distance close to 500 km from the epicenter, shows that the waves that traveled through the uppermost mantle, Pn and Sn, arrive before the direct crustal arrivals, Pg and Sg. The Pn and Sn arrivals are emergent and more difficult to see. Both seismograms are 5 minutes long and are aligned on the first arrival. The vertical scale of both seismograms is the same, with the top one having a peak ground velocity of 1.81 mm/s while the regional recording has a peak ground velocity of 0.40 mm/s, which is almost a five fold decrease in peak ground velocity.

As indicated in Figure 1, for earthquakes recorded at short local distances, the first arrivals are the direct crustal phases Pg and Sg, which propagated through the crust. As the source to receiver distance increases, earthquake recordings may include the Moho-reflected phases PmP and SmS. Beyond the crossover distance, Pn and Sn will be the first arrivals. These waves travel from the source and propagate through the uppermost mantle before turning to the surface again (Storchak et al., 2003). The crossover distance is a function of earthquake depth and crustal thickness, and ranges from 30 km in thin oceanic crust to 200 km in thick continental crust, since crustal thickness can vary from 6 km to 70 km (Mooney et al., 1998). For reference, for a 30 km thick continental crust and assuming typical seismic velocities for the crust and upper mantle, the crossover distance for a surface source would be $\sim$ 150 km.

For most regional earthquake recordings the first arrival is the Pn phase and for S waves, the first arrival is its analog Sn. As seen in Figure 1, the characteristics of the waveforms are different for the local and the regional recordings. The first arrivals Pn and Sn are known to be emergent, compared to the impulsive nature of Pg and Sg. The decay of coda (its envelope) for local recordings tends to follow a one over time pattern (Sato et al., 2012), with the maximum amplitude very close to the first arrival. In contrast, for the regional recording, the envelope of the P and S codas looks more like a spindle, with the maximum amplitudes not as close to the first arrivals. This change in shape is attributed to scattering, which is strongest in the crust and uppermost mantle (Shearer and Earle, 2004). Even though in Figure 1 the secondary S arrival is labeled as Sg, at longer distances, close to 1000 km the high frequency S wave train has been attenuated and only S waves trapped in the crustal waveguide, known as Lg will be the secondary S wave arrival. Lg phases are complex, and can be blocked by changes in crustal structure (Al-Damegh et al., 2004). Careful attention to the demanding task of precise picking of these regional seismic phases leads to improved earthquake catalogs in zones that are otherwise challenging to monitor (Fuenzalida et al., 2013), which suggests that deep-learning-based methods should be extremely useful at these distances.

Figure 2 shows the International Seismological Centre station inventory list with the inverted blue triangles. Seismically active parts of Europe, Japan, New Zealand and the United States, have the greatest concentration of instruments. The green contours and areas are those for which there is a minimum of 5 stations within a radius of 3 degrees. These green outlines enclose those highly instrumented regions of the world where "local" earthquake monitoring with direct crustal

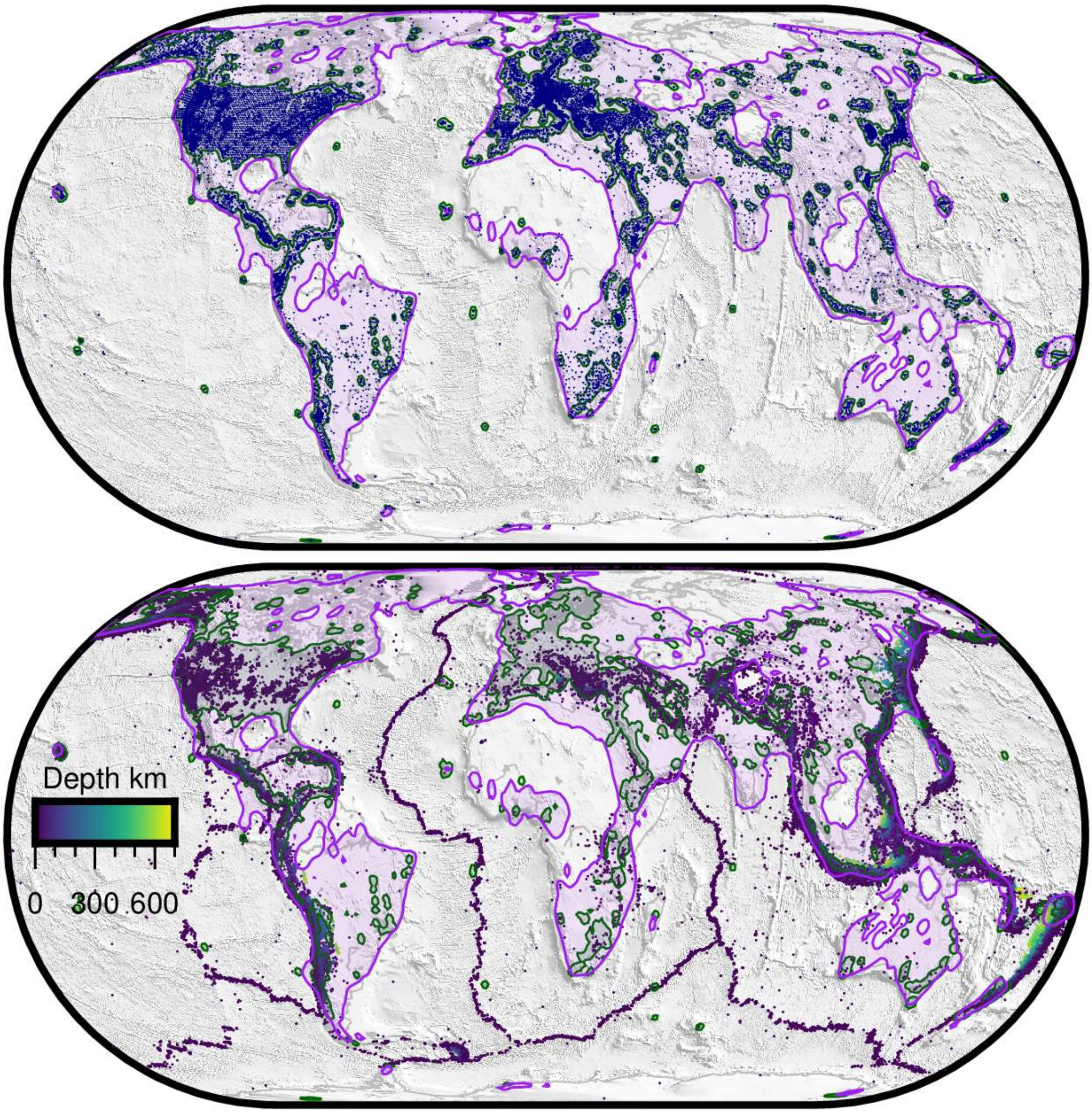


Figure 2 Top. Stations in the ISC inventory list. Bottom. Global seismicity. In both panels the green regions are those for which there are at least 5 stations within a 3 degree radius and the purple regions are those for which there are a minimum of 5 stations within a 10 degree radius and the azimuthal gap for an earthquake within this region would be less than 180 degrees.

phases should be possible. The purple-shaded region indicates where there is a minimum of 5 stations within a 10 degree radius and where the azimuthal gap for an earthquake at each point is less than 180 degrees. These purple regions are those that can be considered suitable for regional monitoring. The area ratio between the green and purple regions is about 10, which indicates that there should be great benefit to more effective regional earthquake monitoring. For the important case of small islands, such as the Azores or Ascension in the South Atlantic, they do not meet the azimuthal gap criteria due to their limited areal footprint. From the point of view of earthquake monitoring, mid-ocean ridges might be considered the least well-monitored seismic zones on Earth. Also, note that this set of stations does not represent current monitoring conditions accurately, since we do not consider information on the lifespan of these stations. For example, the Transportable Array stations across the United States only operated for approximately two years at any particular location, such that much of that area is covered by regional, rather than local, monitoring. That is, monitoring from permanent seismic networks in much of the world is not as effective as this figure suggests.

Adapting the machine learning workflows to regional earthquake monitoring and earthquake catalog building requires adapting the data and the algorithms. The Curated Regional Earthquake Waveforms (CREW) dataset is the first step towards extending deep (i.e., deep learning-based) earthquake catalogs to regional earthquake monitoring, by assembling a high quality benchmark dataset for training deep-learning models.

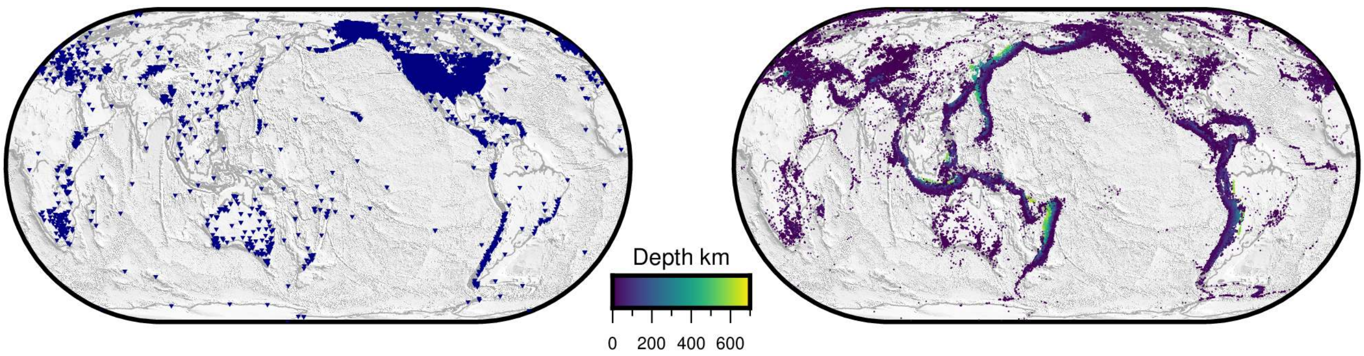


Figure 3 Stations and sources in CREW. Earthquakes are color-coded by depth.

2 Metadata and Data Collection

We queried all datacenters available through Obspy (Beyreuther et al., 2010) to retrieve their earthquake catalogs (see Table 2). We retained only those catalogs that contain information on both P and S arrivals, including phase arrivals P,Pg,Pn and S,Sg,Sn. For instance, catalogs that report only P arrivals were excluded from our workflow. Of the over 30 million metadata entries, we kept only those which for the same station-earthquake pair there were at least one picked arrival of P, Pg, Pn and one picked arrival of S,Sg,Sn on the same trace. That is, we required at least one of the P set and one of the S set to be labeled for each example. The number of traces for which simultaneous P and S information is available is an order of magnitude less than those for which only the first arriving P wave is labeled. Later, we queried all datacenters accesible via Obspy to retrieve the appropriate waveforms in the distance range of 1 to 20 degrees of source to receiver distance, which is the range where the first arrivals are mainly Pn and Sn. Initially, we retrieved 7 minute waveforms, including 2 minutes before the earliest arriving P phase and 5 minutes after. This included all the instruments for each station, encompassing seismometers and accelerometers, with sampling rates ranging from 20 to 200 Hz. We detrended and resampled these data at 100 Hz. We then cut the waveforms randomly so that the earliest arriving P wave is at least 10 seconds after the start of the seismic trace and the total duration is 5 minutes (300 seconds), zero padding when required to complete the 5 minutes. Then, the waveforms were normalized to the absolute peak amplitude among the three channels. Ultimately, we kept the data from the ISC catalog (Storchak et al., 2013) and waveforms from the IRIS DMC (Trabant et al., 2012). The initial pool of data included nearly 3.3 million waveforms and their corresponding arrivals. The sources and receivers represented in the dataset are shown in Figure 3. The database includes 523,294 unique events recorded at 4,071 unique stations around the world. The distribution of earthquakes is representative of global seismology, spanning all latitudes, longitudes and all depths. In contrast, the coverage of receivers is not uniform, as the places with the highest density of instruments are the USA, Chile, and Europe.

Figure 4 displays five examples in the dataset, with their three-component waveforms, along with the arrivals and their labels, and indicate the instrument type and information on the earthquake location and magnitude, as well as the source to receiver distance. These examples are shown for the presence of 2,3 and 4 picked arrivals. Panels A and B represent the most common cases in CREW, where only the first arriving P wave and the first arriving S wave are labeled. For the example in (A), generic P and S labels are provided, whereas for (B), more specific Pn and Sn labels are provided. Panel (C) depicts a case in which three labels are provided, P, Pn and S, but P and Pn represent the same timestamp, so there are effectively two labeled arrivals. (D) shows the case of three distinct phases labeled, Pn, Sn and Sg. The bottom panel of Figure 4 (E) shows an uncommon example, in which the four phases Pn,Pg, Sn and Sg are all labeled, only a few thousand such examples occur in the dataset because most datacenters do not label arrivals other than the first arrival. Note that this example has been bandpass filtered to enhance the visibility of the arrivals. These rare examples typically come from stable continental regions, where the propagation of regional phases is not blocked by crustal and mantle structure (Gök et al., 2003). Panel C is a case in which there are two differently labeled arrivals, Pn and P that are very close in time, corresponding to the same arrival, but having an almost negligible time difference. In cases like this, we preserved all the available labels, but in subsequent workflows we only employed the earliest of the available P arrivals and the earliest of the S arrivals.

3 From Big Data to Good Data

In several fields employing machine learning, performance gains from dataset cleaning and refinement have been shown to surpass those from model architecture improvements (Northcutt et al., 2021a,b). Moreover, if data quality is high, effective training of deep neural networks requires fewer data (Motamedi et al., 2021). This has caused a shift in attention from the quantity of data to the quality of the data, as Data-Centric AI has gained traction (Zha et al., 2023) and led to data-centric initiatives (https://cleanlab.ai/) and competitions (Ng et al., 2021) (https://https-deeplearning-ai.github.io/data-centric-comp/) for more controlled benchmark datasets. (Northcutt et al., 2021a) documented the prevalence of faulty examples for ten of the most used machine learning benchmark datasets including image, text and

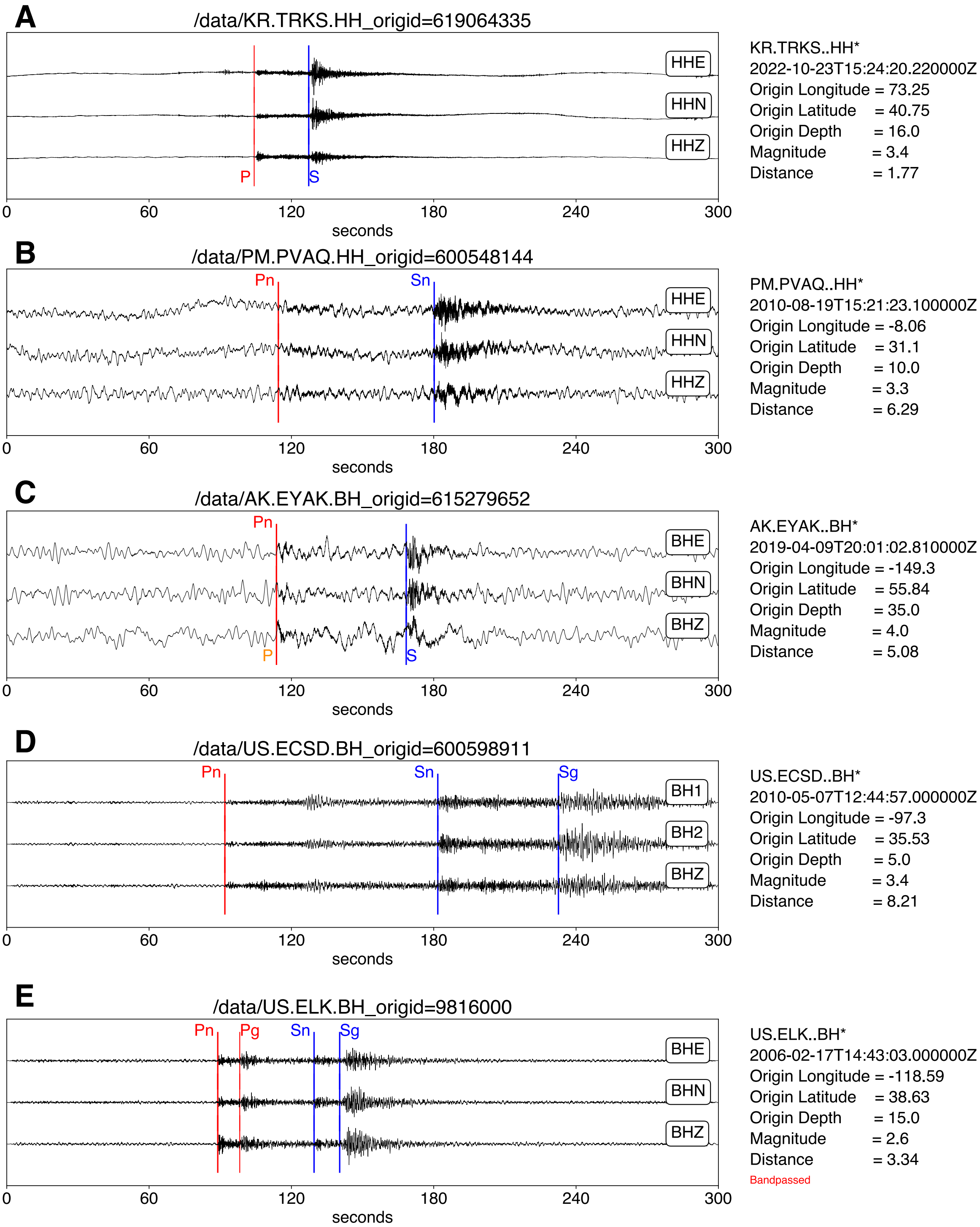


Figure 4 Examples from CREW with 2 (A,B), 3 (B,C), and 4 (E) labeled arrivals. Depth is in km, and distance in degrees. Example in panel (E) has been bandpassed filtered between 1 and 10 Hz to facilitate visualization of the arrivals.

audio, MNIST, CIFAR, and ImageNet among others (https://labelerrors.com/). These examples contain either faulty data or defective or incomplete labels, that end up affecting model selection and performance. Seismological data can contain errors such as inaccurate picked phase arrivals, and seismometer data is prone to corrupted transmission or storage. Strategies to mitigate the effects of bad labels and bad data have been devised (Cordeiro and Carneiro, 2020; Northcutt et al., 2021b), and it remains an active research field. This motivated our shift in approach, from iterating over a fixed dataset and optimizing for model parameters, to fixing the architecture and iteratively improving the dataset by identifying faulty examples, outliers, and edge cases, and/or by synthesizing new examples. CREW was built to have both big and good data.

Upon inspection of the initial dataset, it was clear that there were many faulty examples of various types. We manually checked a random sample of 10,000 examples and classified them into the following categories: (1) good examples, which have seemingly accurate arrival time labels on clean seismograms and (2) bad examples, which have inaccurate arrival time labels, corrupted seismograms, or other problems. Another cate-

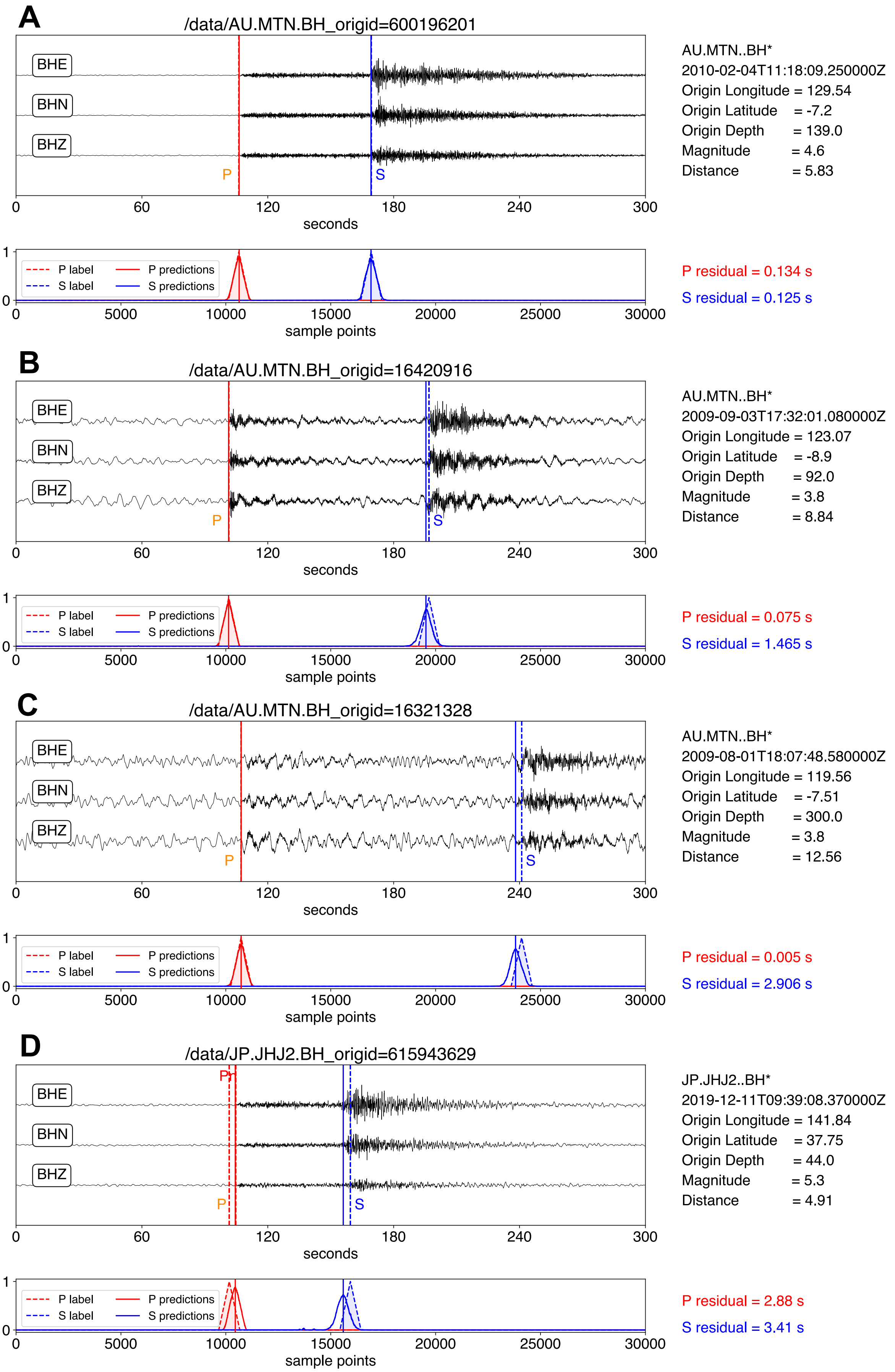


Figure 5 Comparison of labels and predictions, dotted lines are the labels in red for P waves and blue for S waves. The solid lines are the predictions and the inferred picks. The time difference between the two are displayed in the bottom right. Examples in A and B were kept in CREW, while examples in C and D were removed.

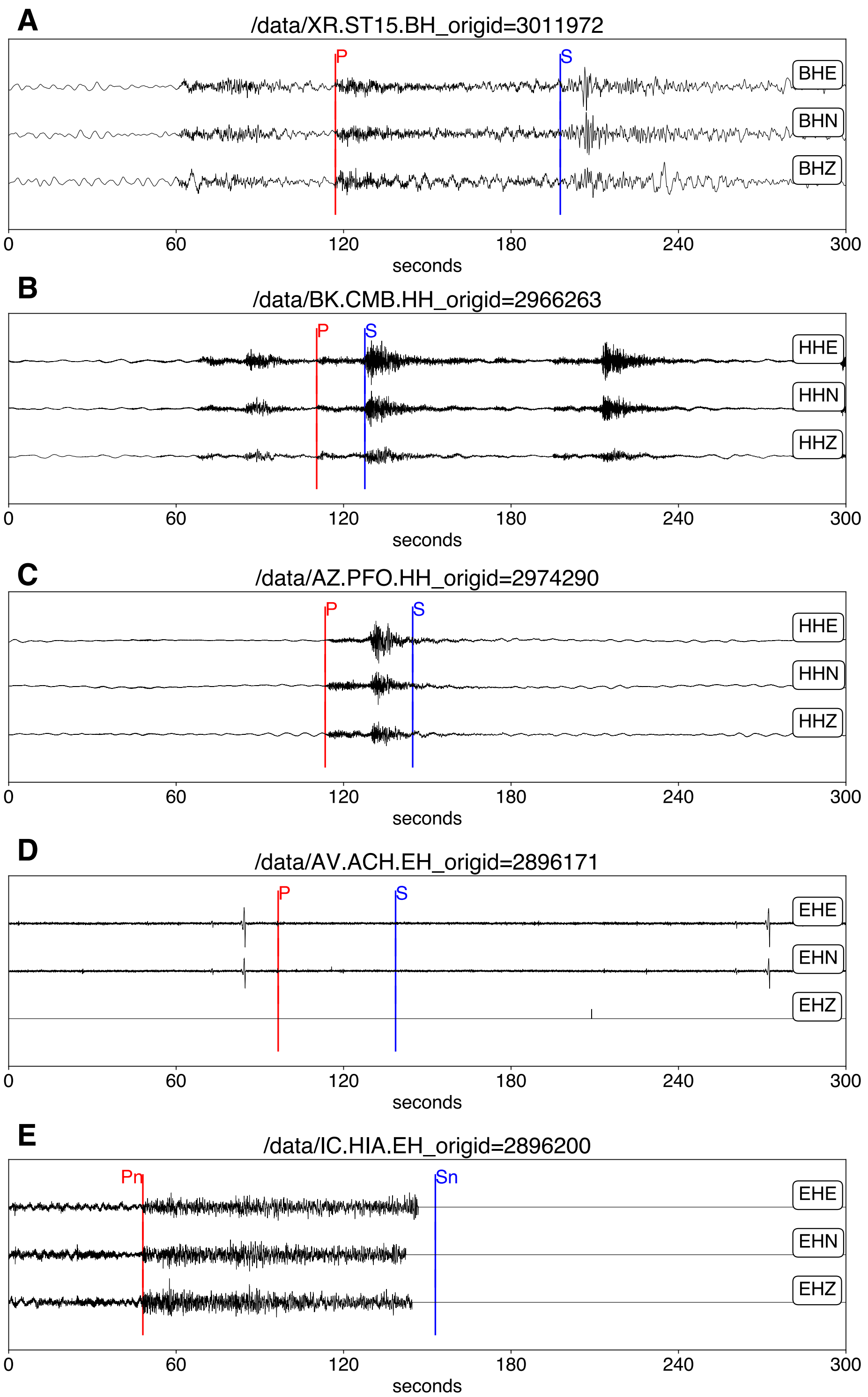


Figure 6 Examples rejected from CREW. (A) uncataloged earthquake. (B) multiple uncataloged earthquakes. (C) accurate P arrival next to an inaccurate S arrival. (D) no earthquake signal visible. (E) accurate Pn arrival but data is incomplete for the Sn arrival.

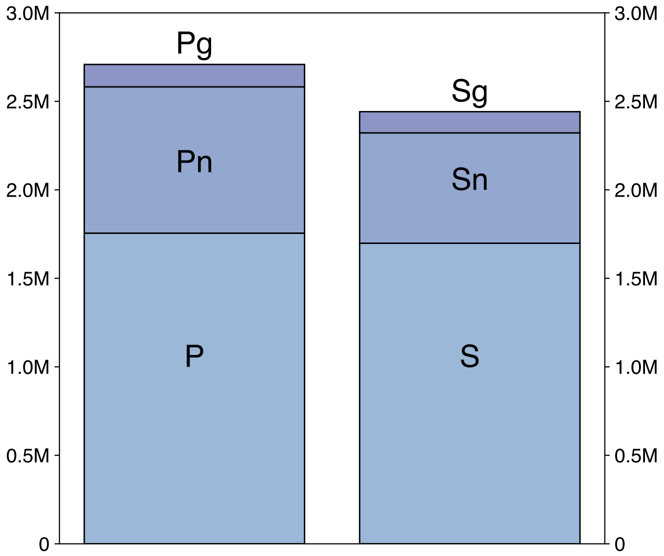


Figure 7 Number of picks in each category.

gory, (3) multiplets, accounted for the class where there are multiple earthquakes, but only one is labeled. From this sorting scheme 72% of the data was deemed as good (category 1) and the remaining 28% was flagged (categories 2 and 3) because it was deemed to contain inaccurate training labels, or corrupted data.

To automate the screening of faulty examples, we trained a convolutional neural network, a U-Net with skip connections, based on the architecture of PhaseNet (Zhu and Beroza, 2019). Our model has several extra layers to process input waveforms that are 10 times longer than those for the original PhaseNet, such that it produces representations in the deepest layers of similar size. The CNN was trained to learn triangular labels that are centered on the pick positions. These labels have a half width of 5 seconds, or 500 sample points on each side, for a total duration of 10 seconds, with a linear increase from 0 to 1 and then a linear decrease from 1 to 0. The labels are the same for both P and S waves, and there is a third channel for noise, which is equal to one minus the label of P and minus the label of S. These triangular labels were made using the earliest arrival among the available arrivals. For instance for the example displayed in Figure 4 the labels used were Pn and Sn. There are multiple examples in the dataset for which multiple arrivals are reported, but in some cases they are very close in time and hard to distinguish. For instance, in Figure 4 panel C there labels for both P and Pn and they are almost overlapping. We leave at the discretion of the user the use of these labels but note that the pruning procedure described here used the earliest arriving among P,Pn, and Pg and the earliest arriving among S,Sn, and Sg. Future research will address working with secondary arrivals.

The training data consists of a mix of data in its raw form, augmented data, and synthetic noise. The details of the architecture, training and deployment of these models and other auxiliary models will be presented in a forthcoming paper. The augmented versions of the data consisted of a superposition of multiple copies of the same example waveforms with a time delay. Depending on the S minus P time, we added a random choice between two or three copies, and with the appropriate labels, those were added to the example. We did this to train the model to work for the frequently encountered scenario where more than one earthquake occurs during a 5-minute window.

Once our phase picker was trained, we applied it to the dataset to remove examples with faulty labels. The criteria used was that the time difference between the dataset labels and the inferred phase picks was under 2 seconds. A large time difference between the label and the prediction was an indication of mistimed arrivals. Figure 5, shows data that passed this criteria and that did not pass it. The delay between labels and predictions are indicated in the bottom right, with red and blue for P and S waves. panel A shows very good alignment of the triangular labels and the model predictions, such that they appear totally superimposed. The predicted arrival times differ by only a tenth of a second, which is an example of what we consider good quality data and labels. Panel B shows good agreement in the P wave, but a delay of nearly a second and a half for the S wave, which is still considered sufficiently good data. Panels C and D show data that was rejected from the dataset because either the P or S prediction differs by more than 2 seconds from the dataset labels. For C, it is the S label that seems to be inaccurate, whereas for D, both the P and S labels are inaccurate, being evident for the S wave, but it requires zooming in to see the P label mislocation.

Figure 6 shows a variety of examples that were flagged as faulty by the described workflow. From top to bottom A through E. (A) Two earthquakes in one window, but only one of them has picked arrivals. (B) At least three uncataloged earthquakes, while only one is labeled. A and B represent the most common way in which the labels are inaccurate. (C) Correct P arrival but faulty S arrival. Without the need for hardcoding a travel time sanity check, our model flagged this type of error. (D) No visible earthquake signal in the waveforms. (E) Data gaps or outages, in this case there is a seemingly accurate Pn arrival, but there is a data gap before the Sn arrival.

The quality-controlled dataset contains 1,599,323 examples (nearly 50% of the initial data pool, nearly 1.1 TB), each a three component waveform sampled at 100 Hz with at least one of P,Pg,Pn and at least one of S,Sg,Sn arrivals. The total number of arrivals is 3,589,986. The proportions of these arrivals are displayed in Figure 7. For the P family there are 1,871,317 arrivals: 1,225,778 for generic P, 564,373 for Pn and 81,166 for Pg. For the S family, there are 1,718,669 arrivals: 1,192,110 for generic S, 446,880 for Sn, and 79,679 for Sg. The relatively low number of Pg and Sg phases is a consequence of excluding data in the 0 to 1 degree distance range. There are multiple existing data sets for those distances as described above.

The resulting dataset consists of 523,294 earthquakes that are globally distributed. The magnitude ranges from 0 to 7.1, with very few earthquakes at either extreme. Figure 8 (left) shows the magnitude-frequency distribution of the earthquakes in CREW as solid bars.

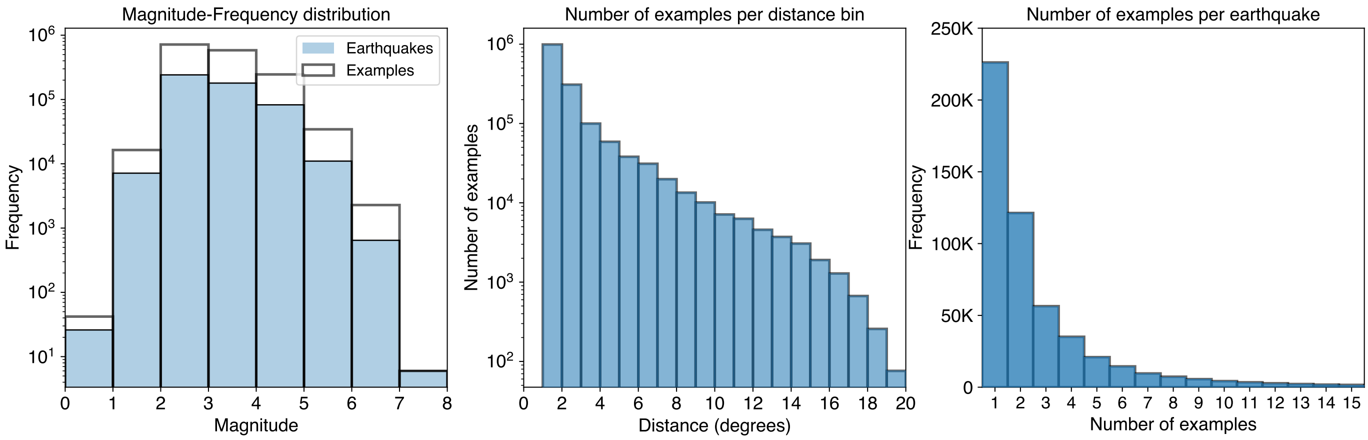


Figure 8 Left. Magnitude-Frequency distribution, the solid bars display the distribution of unique earthquakes in the dataset. The empty bars display the frequency distribution if each example is treated as a separate magnitude. Middle. Distribution of source to receiver distances, which span 1 to 20 degrees. Right. Number of waveforms of examples in the dataset per each unique origin ID, for most earthquakes there is only one or two observations, whereas having more than 5 is rather scarce in the CREW.

The outlined bars show the distribution if each example is treated as a different earthquake, that is counting the same earthquake multiple times. Figure 8 (middle) shows the distance distribution of the examples. Each bar represents a 1 degree distance bin. Most of the data is at the closest distances (distance <4 degrees). The number of recordings decays dramatically with distance, due to the combined effects of amplitude decay and recording limitations, with only larger earthquakes visible at greater distances.

Table 1 summarizes the metadata attributes in CREW. These can be separated into three main categories, station information, earthquake origin information and arrivals information. CREW is stored in hdf5 format, the examples are stored in a group called **data**, where each individual example is named a combination of the station id and the event origin ID. Examples of these names are shown at the top of each panel in Figure 4. For the arrivals, the timestamps are available as well as the sample position corresponding to the location of the arrivals in the waveforms. In Figures 4 and 5 part of the metadata is displayed in the right panels. The names of the variables are mostly in Seisbench format (Woollam et al., 2022), except for the channels list.

The right panel of Figure 8 shows how many examples there are that correspond to a unique earthquake. The most common scenario is that only one recording per earthquake made it through the quality control. Nearly 230,000 earthquakes have only one example, i.e., at least two phase arrivals. On the other hand, 1,251,900 examples correspond to an earthquake with at least 6 phase readings, i.e. a seismic source for which there are at least 3 examples in the dataset. This aspect should be useful for machine learning implementations that perform seismic phase picking incorporating information from multiple stations, e.g. (Feng et al., 2022), or for models that perform earthquake arrival association, e.g. (McBrearty and Beroza, 2023). The plot is clipped at 15, but the earthquake that has the most examples associated with it has 121, which means over 242 picked arrivals. CREW contains more examples with both P and S arrival information than other datasets that cover the same distance range.

We reviewed the data and metadata in CREW, which is global in coverage, containing waveforms from earthquakes from all longitudes, latitudes and depths. CREW includes events up to magnitude 7. Moreover it provides data and labels useful at the single station level as well as the network level, with the majority of the data corresponding to earthquakes with at least 6 arrivals, which should be enough to produce a location.

4 Conclusions and Future Directions

We introduce CREW as a large, high-quality labeled data set for simultaneous regional seismic P and S phase waveforms recorded on seismometers around the world. CREW is the first benchmark data set that focuses on regional phases, rather than phases from local earthquake recordings or teleseismic recordings. Monitoring using regional phases is essential for large parts of the Earth where local monitoring is logistically impractical or is not a high priority due to relatively low seismic hazard. It should also prove useful for the important case of nuclear test ban treaty monitoring. We hope that its availability will enable progress in machine learning for regional earthquake monitoring and structural imaging.

Most machine learning research on seismology has focused on supervised learning (Mousavi and Beroza, 2023), especially for earthquake monitoring, and CREW contributes to this paradigm by curating data with the best available labels for regional first and secondary arrivals. The combination of algorithmic advances and data advances will contribute to multiscale earthquake monitoring.

Future research directions include working on secondary arrivals, such as reflected e.g. PmP, PP or converted waves e.g. SP, PS, that even though not often used for earthquake location, are nevertheless very sensitive to Earth structure and provide insight into the deep interior of the planet. These secondary arrivals are also

Station information	Event Information	Arrivals Information
network_station_code	source_id	{P,Pg,Pn}_arrival_time
station_code	source_origin_time	{P,Pg,Pn}_arrival_sample
channels	source_latitude_deg	{S,Sg,Sn}_arrival_time
station_latitude_deg	source_longitude_deg	{S,Sg,Sn}_arrival_sample
station_longitude_deg	source_depth_km	trace_start_time
station_elevation_m	source_magnitude	
	path_ep_distance_deg	

Table 1 Metadata attributes in CREW. Most ot these attributes are in seisbench convention.

a challenge for machine learning due to the relative scarcity of labeled examples. For these phases, architectures that rely less heavily on labeled data, such as semi-supervised and self-supervised learning that can learn from incomplete labels or partial data might prove successful (Assran et al., 2023). Also, future implementations that aim to characterize the full wavefield by picking all types of seismic phases present should provide improved capabilities for both monitoring and studies of the deep Earth.

Data and Code Availability

CREW is hosted in Stanford University DataFarm: https://redivis.com/datasets/1z6w-e1w70hpmt (https://doi.org/10.57761/60b3-cv76). All codes used to generate and process the dataset are available at https://github.com/albertleonardo/CREW, as well as example data and notebooks. CREW will be made accessible via SeisBench Woollam et al. (2022).

Competing Interests

The authors declare no competing interests. This is a ChatGPT free manuscipt.

Acknowledgements

This work was supported by AFRL under contracts FA9453-19-C-0073 and FA9453-21-9-0054. The authors thank two anonymous reviewers and editors Jack Muir and Théa Ragon for their help in improving this manuscript.

References

Al-Damegh, K., Sandvol, E., Al-Lazki, A., and Barazangi, M. Regional seismic wave propagation (Lg and Sn) and Pn attenuation in the Arabian Plate and surrounding regions. *Geophysical Journal International*, 157(2):775–795, 2004. doi: 10.1111/j.1365-246X.2004.02246.x.

Alaska Earthquake Center, Univ. of Alaska Fairbanks. Alaska Geophysical Network, 1987. doi: 10.7914/SN/AK.

Alaska Volcano Observatory/USGS. Alaska Volcano Observatory, 1988. doi: 10.7914/SN/AV.

Alberta Geological Survey / Alberta Energy Regulator. Regional Alberta Observatory for Earthquake Studies Network, 2013. doi: 10.7914/SN/RV.

Albuquerque Seismological Laboratory (ASL)/USGS. US Geological Survey Networks, 1980. doi: 10.7914/SN/GS.

Albuquerque Seismological Laboratory (ASL)/USGS. China Digital Seismograph Network, 1986. doi: 10.7914/SN/CD.

Albuquerque Seismological Laboratory (ASL)/USGS. United States National Seismic Network, 1990. doi: 10.7914/SN/US.

Albuquerque Seismological Laboratory (ASL)/USGS. New China Digital Seismograph Network, 1992. doi: 10.7914/SN/IC.

Albuquerque Seismological Laboratory (ASL)/USGS. Global Telemetered Seismograph Network (USAF/USGS), 1993. doi: 10.7914/SN/GT.

Albuquerque Seismological Laboratory (ASL)/USGS. New England Seismic Network, 1994. doi: 10.7914/SN/NE.

Albuquerque Seismological Laboratory (ASL)/USGS. Intermountain West Seismic Network, 2003. doi: 10.7914/SN/IW.

Albuquerque Seismological Laboratory/USGS. Central and Eastern US Network, 2013. doi: 10.7914/SN/N4.

Albuquerque Seismological Laboratory/USGS. Global Seismograph Network (GSN - IRIS/USGS), 2014. doi: 10.7914/SN/IU.

Andy Nyblade. Africa Array- Uganda/Tanzania, 2007. doi: 10.7914/SN/ZP_2007.

Andy Nyblade. AfricaArray - Namibia, 2015. doi: 10.7914/SN/8A_2015.

Anne Meltzer. RAMP Virginia, 2011. doi: 10.7914/SN/YC_2011.

Anne Meltzer and Susan Beck. 2016 Pedernales Earthquake Aftershock Deployment Ecuador, 2016. doi: 10.7914/SN/8G_2016.

Anne Sheehan, Francis Wu, and Roger Bilham. Himalayan Seismotectonics, Nepal and Tibet, 2001. doi: 10.7914/SN/YL_2001.

Aristotle University of Thessaloniki. Aristotle University of Thessaloniki Seismological Network, 1981. doi: 10.7914/SN/HT.

Arizona Geological Survey. Arizona Broadband Seismic Network, 2007. doi: 10.7914/SN/AE.

Assran, M., Duval, Q., Misra, I., Bojanowski, P., Vincent, P., Rabbat, M., Lecun, Y., Ballas, N., and Fair, M. A. I. Joint-Embedding Predictive Architecture. *arXiv preprint arXiv:2301.08243v1*, 2023.

Australian National University (ANU, Australia). Australian Seismometers in Schools, 2011. doi: 10.7914/SN/S1.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., and Wassermann, J. ObsPy: A python toolbox for seismology. *Seismological Research Letters*, 81(3):530–533, 2010. doi: 10.1785/gssrl.81.3.530.

Botswana Geoscience Institute. Botswana Seismological Network, 2001. doi: 10.7914/SN/BX.

British Geological Survey. Great Britain Seismograph Network, 1970. doi: 10.7914/AV8J-NC83.

Bureau of Economic Geology, The University of Texas at Austin. Texas Seismological Network, 2016. doi: 10.7914/SN/TX.

California Institute of Technology and United States Geological

Survey Pasadena. Southern California Seismic Network, 1926. doi: 10.7914/SN/CI.

Cascades Volcano Observatory/USGS. Cascade Chain Volcano Monitoring, 2001. doi: 10.7914/SN/CC.

Central Asian Institute for Applied Geosciences. Central Asian Seismic Network of CAIAG, 2008. doi: 10.7914/SN/KC.

Centro de Geociencias, UNAM. Permanent Seismic Network of Queretaro State, Mexico, 2003. doi: 10.7914/SN/MG.

Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), Ensenada. Red Sísmica del Noroeste de México, 1980. doi: 10.7914/SN/BC.

Charles University in Prague (Czech), Institute of Geonics, Institute of Geophysics, Academy of Sciences of the Czech Republic, Institute of Physics of the Earth Masaryk University (Czech), and Institute of Rock Structure and Mechanics. Czech Regional Seismic Network, 1973. doi: 10.7914/SN/CZ.

Cindy Ebinger. AFAR07, 2007. doi: 10.7914/SN/ZE_2007.

Cole, H. M., Yeck, W. L., and Benz, H. M. MLAAPDE: A Machine Learning Dataset for Determining Global Earthquake Source Parameters. *Seismological Research Letters*, 20(20):1–11, 2023. doi: 10.1785/0220230021.

Colorado Geological Survey. Colorado Geological Survey Seismic Network, 2016. doi: 10.7914/SN/C0.

Cordeiro, F. R. and Carneiro, G. A Survey on Deep Learning with Noisy Labels: How to train your model when you cannot trust on the annotations? *Proceedings - 2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images, SIBGRAPI 2020*, pages 9–16, 2020. doi: 10.1109/SIBGRAPI51738.2020.00010.

Cynthia Ebinger. Magadi-Natron Magmatic Rifting Studies, 2013. doi: 10.7914/SN/XJ_2013.

DANA . Dense Array for North Anatolia, 2012. doi: 10.7914/SN/YH_2012.

Douglas Christensen, Roger Hansen, and Geoff Abers. Broadband Experiment Across the Alaska Range, 1999. doi: 10.7914/SN/XE_1999.

Douglas Wiens. Southwest Pacific Seismic Experiment, 1993. doi: 10.7914/SN/XB_1993.

Douglas Wiens. Mantle serpentinization and water cycling through the Mariana Trench and Forearc, 2012. doi: 10.7914/SN/XF_2012.

Dublin Institute for Advanced Studies. Irish National Seismic Network, 1993. doi: 10.7914/SN/EI.

Federal Institute for Geosciences and Natural Resources. German Regional Seismic Network (GRSN), 1976. doi: 10.25928/MBX6-HR74.

Feng, T., Mohanna, S., and Meng, L. EdgePhase: A Deep Learning Model for Multi-Station Seismic Phase Picking. *Geochemistry, Geophysics, Geosystems*, 23(11), 2022. doi: 10.1029/2022GC010453.

Frank Vernon. ANZA Regional Network, 1982. doi: 10.7914/SN/AZ.

Frank Vernon. Broadband Seismic Characterization of the Arabian Shield, 1995. doi: 10.7914/SN/XI_1995.

Fuenzalida, A., Schurr, B., Lancieri, M., Sobiesiak, M., and Madariaga, R. High-resolution relocation and mechanism of aftershocks of the 2007 Tocopilla (Chile) earthquake. *Geophysical Journal International*, 194(2):1216–1228, 2013. doi: 10.1093/gji/ggt163.

Geoffrey A. Abers and Karen M. Fischer. Broadband Tomography Under Costa Rica and Nicaragua, 2003. doi: 10.7914/SN/YO_2003.

GEOFON Data Centre. GEOFON Seismic Network, 1993. doi: 10.14470/TR560404.

Geological Survey of Canada. Public Safety Geoscience Program Canadian Research Network, 2013. doi: 10.7914/SN/PQ.

Geophysical Survey of the National Academy of sciences of Tajikistan. Tajikistan National Seismic Network, 2009. doi: 10.7914/SN/TJ.

Geoscience Australia. Australian National Seismograph Network Data Collection, 2021. doi: 10.26186/144675.

GNS Science. GeoNet Aotearoa New Zealand Seismic Digital Waveform Dataset, 2021. doi: 10.21420/G19Y-9D40.

Gök, R., Sandvol, E., Türkelli, N., Seber, D., and Barazangi, M. Sn attenuation in the Anatolian and Iranian plateau and surrounding regions. *Geophysical Research Letters*, 30(24):3–6, 2003. doi: 10.1029/2003GL018020.

Idaho National Laboratory. INL Seismic Monitoring Program, 1972. doi: 10.7914/SN/IE.

Incorporated Research Institutions For Seismology. SINGLE STATION, 1970. doi: 10.7914/SN/SS.

Institut Cartogràfic i Geològic de Catalunya. Catalan Seismic Network, 1984. doi: 10.7914/SN/CA.

Institut De Physique Du Globe De Paris (IPGP). GNSS, seismic broadband and strong motion permanent networks in West Indies, 2008. doi: 10.18715/ANTILLES.WI.

Institut de physique du globe de Paris (IPGP) and École et Observatoire des Sciences de la Terre de Strasbourg (EOST). GEOSCOPE, French Global Network of broad band seismic stations, 1982. doi: 10.18715/GEOSCOPE.G.

Institute of Earth Sciences, Academia Sinica, Taiwan. Broadband Array in Taiwan for Seismology, 1996. doi: 10.7914/SN/TW.

Institute of Geophysics China Earthquake Administration (IGPCEA). China National Seismic Network, Data Management Centre of China National Seismic Network at Institute of Geophysics, CEA, 2000. doi: 10.7914/SN/CB.

Institute of GeoSciences (IGEO), Polytechnic University of Tirana (PUT). Albanian Seismological Network, 2002. doi: 10.7914/SN/AC.

Instituto Dom Luiz - Faculdade de Ciências da Universidade de Lisboa. Instituto Dom Luiz (IDL) - Faculdade de Ciencias Universidade de Lisboa, 2003. doi: 10.7914/SN/LX.

Instituto Nacional de Sismologia, Vulcanologia, Meteorologia e Hidrologia (INSIVUMEH). Red Sismologica Nacional, 1976. doi: 10.7914/SN/GI.

Instituto Nicaraguense de Estudios Territoriales (INETER). Nicaraguan Seismic Network, 1975. doi: 10.7914/SN/NU.

Instituto Politecnico Loyola. Observatorio Sismológico Politécnico Loyola, 2012. doi: 10.7914/SN/LO.

Instituto Português do Mar e da Atmosfera, I.P. Portuguese National Seismic Network, 2006. doi: 10.7914/SN/PM.

IRIS Transportable Array. USArray Transportable Array, 2003. doi: 10.7914/SN/TA.

Istituto Nazionale di Oceanografia e di Geofisica Sperimentale. Antarctic Seismographic Argentinean Italian Network - ASAIN, 1992. doi: 10.7914/SN/AI.

Istituto Nazionale di Oceanografia e di Geofisica Sperimentale. International OGS Network, 2014. doi: 10.7914/SN/IO.

Istituto Nazionale di Oceanografia e di Geofisica Sperimentale - OGS. North-East Italy Seismic Network, 2016. doi: 10.7914/SN/OX.

Jay Pulliam. Greater Antilles Seismic Program, 2013. doi: 10.7914/SN/ZC_2013.

Jim Gaherty, Cindy Ebinger, Andy Nyblade, and Donna Shillington. Study of Extension and maGmatism in Malawi aNd Tanzania, 2013. doi: 10.7914/SN/YQ_2013.

Jim Ni, Tom Hearn, and John Nabelek. INDEPTH-III, 1997. doi: 10.7914/SN/XR_1997.

Kandilli Observatory And Earthquake Research Institute, Boğaziçi University. Kandilli Observatory And Earthquake Research Institute (KOERI), 1971. doi: 10.7914/SN/KO.

Kate Miller. Bhutan Pilot Experiment, 2002. doi: 10.7914/SN/XA_2002.

Kentucky Geological Survey/Univ. of Kentucky. Kentucky Seismic and Strong Motion Network, 1982. doi: 10.7914/SN/KY.

KNDC/Institute of Geophysical Research (Kazakhstan). Kazakhstan Network, 1994. doi: 10.7914/SN/KZ.

KNMI. Netherlands Seismic and Acoustic Network, 1993. doi: 10.21944/E970FD34-23B9-3411-B366-E4F72877D2C5.

Kyrgyz Institute of Seismology, IVTAN/KIS and University of California, San Diego. Kyrgyz Seismic Telemetry Network, 1991. doi: 10.7914/SN/KN.

Kyrgyz Institute of Seismology, KIS . Kyrgyz Digital Network, 2007. doi: 10.7914/SN/KR.

Lamont Doherty Earth Observatory (LDEO), Columbia University. Lamont-Doherty Cooperative Seismographic Network, 1970. doi: 10.7914/SN/LD.

LTD Seismological Experience and Methodology Expedition of the Committee of Science of the Ministry of Education and Science of the Republic of Kazakhstan. Seismic network of the Seismological Experience and Methodology Expedition CS MES RK, 2003. doi: 10.7914/SN/QZ.

Magrini, F., Jozinović, D., Cammarano, F., Michelini, A., and Boschi, L. Local earthquakes detection: A benchmark dataset of 3-component seismograms built on a global scale. *Artificial Intelligence in Geosciences*, 1(July):1–10, 2020. doi: 10.1016/j.aiig.2020.04.001.

McBrearty, I. W. and Beroza, G. C. Earthquake Phase Association with Graph Neural Networks. *Bulletin of the Seismological Society of America*, 113(2):524–547, 2023. doi: 10.1785/0120220182.

MedNet Project Partner Institutions. Mediterranean Very Broadband Seismographic Network (MedNet), 1990. doi: 10.13127/SD/FBBBTDTD6Q.

Michael West. The Colima Deep Seismic Experiment: Imaging the Magmatic Root of Colima Volcano, 2006. doi: 10.7914/SN/ZA_2006.

Michelini, A., Cianetti, S., Gaviano, S., Giunchi, C., Jozinović, D., and Lauciani, V. INSTANCE-the Italian seismic dataset for machine learning. *Earth System Science Data*, 13(12):5509–5544, 2021. doi: 10.5194/essd-13-5509-2021.

Montana Bureau of Mines and Geology/Montana Tech (MBMG, MT USA). Montana Regional Seismic Network, 1982. doi: 10.7914/SN/MB.

Mooney, W. D., Laske, G., and Masters, T. G. CRUST 5.1: A global crustal model at 5° × 5°. *Journal of Geophysical Research: Solid Earth*, 103(1):727–747, 1998. doi: 10.1029/97jb02122.

Motamedi, M., Sakharnykh, N., and Kaldewey, T. A Data-Centric Approach for Training Deep Neural Networks with Less Data. *35th Conference on Neural Information Processing Systems (NeurIPS)*, 2021. doi: 10.48550/arXiv.2110.03613.

Mousavi, S. M. and Beroza, G. C. Machine Learning in Earthquake Seismology. *Annual Review of Earth and Planetary Sciences*, 51 (1), 2023. doi: 10.1146/annurev-earth-071822-100323.

Mousavi, S. M., Sheng, Y., Zhu, W., and Beroza, G. C. STanford EArthquake Dataset (STEAD): A Global Data Set of Seismic Signals for AI. *IEEE Access*, 7:179464–179476, 2019. doi: 10.1109/ACCESS.2019.2947848.

Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., and Beroza, G. C. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature Communications*, 11(1):1–12, 2020. doi: 10.1038/s41467-020-17591-w.

Mozambique Rift Tomography, George Helffrich, and Fonseca, J F B D. Mozambique Rift Tomography, 2011. doi: 10.7914/SN/6H_2011.

National Centre for Seismological Research (CENAIS Cuba). Servicio Sismologico Nacional de Cuba, 1998. doi: 10.7914/SN/CW.

National Observatory of Athens, Institute of Geodynamics, Athens. National Observatory of Athens Seismic Network, 1975. doi: 10.7914/SN/HL.

National Seismological Centre. Nepal Kathmandu, 1978. doi: 10.7914/SN/NK.

National Seismological Centre. Centro Nacional de Sismologia, 1998. doi: 10.7914/SN/DR.

Natural Resources Canada. Canadian National Seismograph Network, 1975. doi: 10.7914/SN/CN.

Ng, A., Laird, D., and He, L. Data-Centric AI Competition. *https://https-deeplearning-ai.github.io/data-centric-comp/*, 2021. https://https-deeplearning-ai.github.io/data-centric-comp/.

Ni, Y., Hutko, A. R., Skene, F., Denolle, M. A., Malone, S. D., Bodin, P., Hartog, J. R., and Wright, A. K. Curated Pacific Northwest AI-ready Seismic Dataset. *Seismica*, 2, 2023. doi: 10.26443/seismica.v2i1.368.

NOAA National Oceanic and Atmospheric Administration (USA). National Tsunami Warning Center Alaska Seismic Network, 1967. doi: 10.7914/SN/AT.

Norsar. NORSAR Station Network, 1971. doi: 10.21348/D.NO.0001.

Northcutt, C. G., Athalye, A., and Mueller, J. Pervasive Label Errors in Test Sets Destabilize Machine Learning Benchmarks. *arXiv preprint arXiv:2103.14749*, pages 1–16, 2021a. http://arxiv.org/abs/2103.14749.

Northcutt, C. G., Jiang, L., and Chuang, I. L. Confident learning: Estimating uncertainty in dataset labels. *Journal of Artificial Intelligence Research*, 70:1373–1411, 2021b. doi: 10.1613/JAIR.1.12125.

Northern California Earthquake Data Center. Berkeley Digital Seismic Network (BDSN), 2014. doi: 10.7932/BDSN.

Obsercatorio Vulcanológico y Sismológico de Costa Rica, Universidad Nacional. Observatorio Vulcanológico y Sismológico de Costa Rica, 1984. doi: 10.7914/SN/OV.

Observatório Nacional, Rio de Janeiro, RJ. Rede Sismográfica do Sul e do Sudeste, 2011. doi: 10.7914/SN/ON.

Ocean Networks Canada. NEPTUNE seismic stations, 2009. doi: 10.7914/SN/NV.

OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale) and University of Trieste. North-East Italy Broadband Network, 2002. doi: 10.7914/SN/NI.

Ohio Geological Survey. Ohio Seismic Network, 1999. doi: 10.7914/SN/OH.

Oklahoma Geological Survey. Oklahoma Seismic Network, 1978. doi: 10.7914/SN/OK.

Oklahoma Geological Survey. Oklahoma Consolidated Temporary Seismic Networks, 2018. doi: 10.7914/SN/O2.

Pacific Tsunami Warning Center. Pacific Tsunami Warning Seismic System, 1965. doi: 10.7914/SN/PT.

Paul Silver. Anatomy of an Archean Craton, South Africa, A Multidisciplinary Experiment across the Kaapvaal Craton, 1997. doi: 10.7914/SN/XA_1997.

Penn State University. AfricaArray, 2004a. doi: 10.7914/SN/AF.

Penn State University. Pennsylvania State Seismic Network, 2004b. doi: 10.7914/SN/PE.

Red Sismica Volcan Baru. ChiriNet, 2000. doi: 10.7914/SN/PA.

Regional Integrated Multi-Hazard Early Warning System (RIMES Thailand). Regional Integrated Multi-Hazard Early Warning System, 2008. doi: 10.7914/SN/RM.

Roger Hansen and Gary Pavlis. Collaborative Research: St. Elias Erosion/Tectonics Project, 2005. doi: 10.7914/SN/XZ_2005.

Ross, Z. E., Meier, M. A., Hauksson, E., and Heaton, T. H. Generalized seismic phase detection with deep learning. *Bulletin of the Seismological Society of America*, 108(5):2894–2901, 2018. doi: 10.1785/0120180080.

Royal Observatory of Belgium. Belgian Seismic Network, 1985. doi: 10.7914/SN/BE.

San Fernando Royal Naval Observatory (ROA), Universidad Complutense De Madrid (UCM), Helmholtz-Zentrum Potsdam Deutsches GeoForschungsZentrum (GFZ), Universidade De Évora (UEVORA, Portugal), and Institute Scientifique Of Rabat (ISRABAT, Morocco). The Western Mediterranean BB seismic Network, 1996. doi: 10.14470/JZ581150.

Sato, H., Fehler, M. C., and Maeda, T. *Seismic Wave Propagation and Scattering in the Heterogeneous Earth : Second Edition*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012. doi: 10.1007/978-3-642-23029-5.

Scripps Institution of Oceanography. Global Seismograph Network - IRIS/IDA, 1986. doi: 10.7914/SN/II.

Seismological Laboratory of University of Basrah. Iraqi Seismic Observatory, 2014. doi: 10.7914/SN/MP.

Servicio Geológico Colombiano. Red Sismologica Nacional de Colombia, 1993. doi: 10.7914/SN/CM.

Shearer, P. M. and Earle, P. S. The global short-period wavefield modelled with a Monte Carlo seismic phonon method. *Geophysical Journal International*, 158(3):1103–1117, 2004. doi: 10.1111/j.1365-246X.2004.02378.x.

Stephane Rondenay. Multi-disciplinary Experiments for Dynamic Understanding of Subduction under the Aegean Sea, 2006. doi: 10.7914/SN/XS_2006.

Steve Roecker and Ray Russo. RAMP response for 2010 Chile earthquake, 2010. doi: 10.7914/SN/XY_2010.

Storchak, D. A., Schweitzer, J., and Bormann, P. The IASPEI standard seismic phase list. *Seismological Research Letters*, 74(6): 761–772, 2003. doi: 10.1785/gssrl.74.6.761.

Storchak, D. A., Giacomo, D. D., Bondár, I., Engdahl, E. R., Harris, J., Lee, W. H., Villaseñor, A., and Bormann, P. Public release of the ISC-GEM global instrumental earthquake catalogue (1900-2009). *Seismological Research Letters*, 84(5):810–815, 2013. doi: 10.1785/0220130034.

Susan Beck, Terry Wallace, and George Zandt. Slab Geometry in the Southern Andes, 2000. doi: 10.7914/SN/YC_2000.

Susan Schwartz. Imaging the Seismogenic Zone with Geodesy and Seismology, 1999. doi: 10.7914/SN/XY_1999.

Swiss Seismological Service (SED) At ETH Zurich. National Seismic Networks of Switzerland, 1983. doi: 10.12686/SED/NETWORKS/CH.

Sylvie Leroy, Derek Keir, and Graham Stuart. Young Conjugate Margins Lab in the Gulf of Aden, 2009. doi: 10.7914/SN/XW_2009.

Texas Seismological Network, Alexandros Savvaidis. Texas Seismological Network, 2018. doi: 10.7914/SN/4T_2018.

Trabant, C., Hutko, A. R., Bahavar, M., Karstens, R., Ahern, T., and Aster, R. Data products at the IRIS DMC: Stepping stones for research and other applications. *Seismological Research Letters*, 83(5):846–854, 2012. doi: 10.1785/0220120032.

UC Santa Barbara. UC Santa Barbara Engineering Seismology Network, 1989. doi: 10.7914/SN/SB.

Universidad de Chile. Red Sismologica Nacional, 2012. doi: 10.7914/SN/C1.

Universidad de Costa Rica. Red Sismológica Nacional de Costa Rica (RSN: UCR-ICE), 2016. doi: 10.15517/TC.

Universidad Nacional Autónoma de México (UNAM). Red Sísmica Mexicana, 1970. doi: 10.21766/SSNMX/SN/MX.

University of Nevada, Reno. Nevada Seismic Network, 1971. doi: 10.7914/SN/NN.

University of Oregon. Pacific Northwest Seismic Network - University of Oregon, 1990. doi: 10.7914/SN/UO.

University of Ottawa (uOttawa Canada). Yukon-Northwest Seismic Network, 2013. doi: 10.7914/SN/NY.

University of Puerto Rico. Puerto Rico Seismic Network & Puerto Rico Strong Motion Program, 1986. doi: 10.7914/SN/PR.

University of South Carolina. South Carolina Seismic Network, 1987. doi: 10.7914/SN/CO.

University of Utah. University of Utah Regional Seismic Network, 1962. doi: 10.7914/SN/UU.

University of Utah. Yellowstone National Park Seismograph Network, 1983. doi: 10.7914/SN/WY.

University of Washington. Pacific Northwest Seismic Network - University of Washington, 1963. doi: 10.7914/SN/UW.

USGS Alaska Anchorage. USGS Northern Mariana Islands Network, 2000. doi: 10.7914/SN/MI.

USGS Hawaiian Volcano Observatory (HVO). Hawaiian Volcano Observatory Network, 1956. doi: 10.7914/SN/HV.

Utrecht University (UU Netherlands). NARS, 1983. doi: 10.7914/SN/NR.

Vadim Levin. Retreating-Trench, Extension, and Accretion LTectonics: A Multidisciplinary Study of the Northern Apennines, 2003. doi: 10.7914/SN/YI_2003.

Various Institutions. International Miscellaneous Stations, 1965. doi: 10.7914/VEFQ-VH75.

Wenyuan Fan, Yihe Huang, and S. Shawn Wei. Lake-induced earthquakes in Lake Erie, 2018. doi: 10.7914/SN/7D_2018.

Won Sang Lee and Yongcheol Park. Korea Polar Observation Network, 2013. doi: 10.7914/SN/KP.

Woollam, J., Rietbrock, A., Bueno, A., and De Angelis, S. Convolutional neural network for seismic phase classification, performance demonstration over a local seismic network. *Seismological Research Letters*, 90(2 A):491–502, 2019. doi: 10.1785/0220180312.

Woollam, J., Münchmeyer, J., Tilmann, F., Rietbrock, A., Lange, D., Bornstein, T., Diehl, T., Giunchi, C., Haslinger, F., Jozinović, D., Michelini, A., Saul, J., and Soto, H. SeisBench-A Toolbox for Machine Learning in Seismology. *Seismological Research Letters*, 93(3):1695–1709, 2022. doi: 10.1785/0220210324.

Yeck, W. L., Patton, J. M., Ross, Z. E., Hayes, G. P., Guy, M. R., Nick, B., Shelly, D. R., Benz, H. M., and Earle, P. S. Leveraging Deep Learning in Global 24 / 7 Real-Time Earthquake Monitoring at the National Earthquake Information Center. *Seismological Research Letters*, 2020. doi: 10.1785/0220200178.Supplemental.

ZAMG - Zentralanstalt für Meterologie und Geodynamik. Austrian Seismic Network, 1987. doi: 10.7914/SN/OE.

Zha, D., Bhat, Z. P., Lai, K.-H., Yang, F., and Hu, X. Data-centric AI:

Perspectives and Challenges. *arXiv preprint arXiv:2301.04819*, 2023. http://arxiv.org/abs/2301.04819.

Zhu, W. and Beroza, G. C. PhaseNet: A deep-neural-network-based seismic arrival-time picking method. *Geophysical Journal International*, 216(1):261–273, 2019. doi: 10.1093/gji/ggy423.

Table 2 Seismological Networks used in CREW

4T	Texas Seismological Network, Alexandros Savvaidis (2018)
6H	Mozambique Rift Tomography et al. (2011)
7D	Wenyuan Fan et al. (2018)
7F	Central Arkansas Induced Eathquakes 2010-2011
8A	Andy Nyblade (2015)
8G	Anne Meltzer and Susan Beck (2016)
AC	Institute of GeoSciences (IGEO), Polytechnic University of Tirana (PUT) (2002)
AE	Arizona Geological Survey (2007)
AG	Arkansas Seismic Network
AF	Penn State University (2004a)
AI	Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (1992)
AK	Alaska Earthquake Center, Univ. of Alaska Fairbanks (1987)
AR	Northern Arizona Network
AT	NOAA National Oceanic and Atmospheric Administration (USA) (1967)
AU	Geoscience Australia (2021)
AV	Alaska Volcano Observatory/USGS (1988)
AY	Haitian Seismic Network
AZ	Frank Vernon (1982)
BC	Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), Ensenada (1980)
BE	Royal Observatory of Belgium (1985)
BK	Northern California Earthquake Data Center (2014)
BL	Brazilian Lithospheric Seismic Project
BX	Botswana Geoscience Institute (2001)
C	Chilean National Seismic Network
C0	Colorado Geological Survey (2016)
C1	Universidad de Chile (2012)
C8	Canadian Seismic Research Network
CA	Institut Cartogràfic i Geològic de Catalunya (1984)
CB	Institute of Geophysics China Earthquake Administration (IGPCEA) (2000)
CC	Cascades Volcano Observatory/USGS (2001)
CD	Albuquerque Seismological Laboratory (ASL)/USGS (1986)
CH	Swiss Seismological Service (SED) At ETH Zurich (1983)
CI	California Institute of Technology and United States Geological Survey Pasadena (1926)
CK	CAREMON Central Asian Cross-border network
CM	Servicio Geológico Colombiano (1993)
CN	Natural Resources Canada (1975)
CO	University of South Carolina (1987)
CS	Caucusus Array (CS)
CW	National Centre for Seismological Research (CENAIS Cuba) (1998)
CY	Cayman Islands Seismic Network
CZ	Charles University in Prague (Czech) et al. (1973)
DK	Danish Seismological Network
DR	National Seismological Centre (1998)
EC	Ecuador Seismic Network
EI	Dublin Institute for Advanced Studies (1993)
ET	CERI Southern Appalachian Seismic Network

Table 2 continued: Seismological Networks used in CREW

G	Institut de physique du globe de Paris (IPGP) and École et Observatoire des Sciences de la Terre de Strasbourg (EOST) (1982)
GB	British Geological Survey (1970)
GE	GEOFON Data Centre (1993)
GI	Instituto Nacional de Sismologia, Vulcanologia, Meteorologia e Hidrologia (INSIVUMEH) (1976)
GO	National Seismic Network of Georgia
GR	Federal Institute for Geosciences and Natural Resources (1976)
GS	Albuquerque Seismological Laboratory (ASL)/USGS (1980)
GT	Albuquerque Seismological Laboratory (ASL)/USGS (1993)
HK	Hong Kong Seismograph Network
HL	National Observatory of Athens, Institute of Geodynamics, Athens (1975)
HT	Aristotle University of Thessaloniki (1981)
HV	USGS Hawaiian Volcano Observatory (HVO) (1956)
IC	Albuquerque Seismological Laboratory (ASL)/USGS (1992)
IE	Idaho National Laboratory (1972)
II	Scripps Institution of Oceanography (1986)
IM	Various Institutions (1965)
IN	National Seismic Network of India
IO	Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (2014)
IU	Albuquerque Seismological Laboratory/USGS (2014)
IW	Albuquerque Seismological Laboratory (ASL)/USGS (2003)
JP	Japan Meteorological Agency Seismic Network
KC	Central Asian Institute for Applied Geosciences (2008)
KG	Korean Seismic Network - KIGAM
KN	Kyrgyz Institute of Seismology, IVTAN/KIS and University of California, San Diego (1991)
KO	Kandilli Observatory And Earthquake Research Institute, Boğaziçi University (1971)
KP	Won Sang Lee and Yongcheol Park (2013)
KR	Kyrgyz Institute of Seismology, KIS (2007)
KS	Korea National Seismography Network (KNSN-KMA) (KNSN)
KY	Kentucky Geological Survey/Univ. of Kentucky (1982)
KZ	KNDC/Institute of Geophysical Research (Kazakhstan) (1994)
LB	Leo Brady Network (LB)
LD	Lamont Doherty Earth Observatory (LDEO), Columbia University (1970)
LO	Instituto Politecnico Loyola (2012)
LX	Instituto Dom Luiz - Faculdade de Ciências da Universidade de Lisboa (2003)
MB	Montana Bureau of Mines and Geology/Montana Tech (MBMG, MT USA) (1982)
MG	Centro de Geociencias, UNAM (2003)
MI	USGS Alaska Anchorage (2000)
MN	MedNet Project Partner Institutions (1990)
MP	Seismological Laboratory of University of Basrah (2014)
MX	Universidad Nacional Autónoma de México (UNAM) (1970)
MY	Malaysian National Seismic Network
N4	Albuquerque Seismological Laboratory/USGS (2013)
NE	Albuquerque Seismological Laboratory (ASL)/USGS (1994)
NI	OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale) and University of Trieste (2002)
NK	National Seismological Centre (1978)
NL	KNMI (1993)
NM	Cooperative New Madrid Seismic Network
NN	University of Nevada, Reno (1971)
NO	Norsar (1971)
NR	Utrecht University (UU Netherlands) (1983)
NU	Instituto Nicaraguense de Estudios Territoriales (INETER) (1975)
NV	Ocean Networks Canada (2009)
NY	University of Ottawa (uOttawa Canada) (2013)

Table 2 continued: Seismological Networks used in CREW

NZ	GNS Science (2021)
O2	Oklahoma Geological Survey (2018)
OC	Observatorio Sismológico CIGEOBIO CONICET (OSCO)
OE	ZAMG - Zentralanstalt für Meterologie und Geodynamik (1987)
OH	Ohio Geological Survey (1999)
OK	Oklahoma Geological Survey (1978)
ON	Observatório Nacional, Rio de Janeiro, RJ (2011)
OV	Obsercatorio Vulcanológico y Sismológico de Costa Rica, Universidad Nacional (1984)
OX	Istituto Nazionale di Oceanografia e di Geofisica Sperimentale - OGS (2016)
PA	Red Sismica Volcan Baru (2000)
PE	Penn State University (2004b)
PL	Polish Seismological Network
PM	Instituto Português do Mar e da Atmosfera, I.P. (2006)
PO	Portable Observatories for Lithospheric Analysis and Research Investigating Seismicity (POLARIS)
PQ	Geological Survey of Canada (2013)
PR	University of Puerto Rico (1986)
PS	Pacific21 (ERI/STA)
PT	Pacific Tsunami Warning Center (1965)
QZ	LTD Seismological Experience and Methodology Expedition of the Committee of Science of the Ministry of Education and Science of the Republic of Kazakhstan (2003)
RM	Regional Integrated Multi-Hazard Early Warning System (RIMES Thailand) (2008)
RV	Alberta Geological Survey / Alberta Energy Regulator (2013)
S1	Australian National University (ANU, Australia) (2011)
SB	UC Santa Barbara (1989)
SC	New Mexico Tech Seismic Network
SE	Southeastern Appalachian Cooperative Seismic Network
SS	Incorporated Research Institutions For Seismology (1970)
SV	Servicio Nacional de Estudios Territoriales (SNET), El Salvador (SNET-BB)
TA	IRIS Transportable Array (2003)
TC	Universidad de Costa Rica (2016)
TJ	Geophysical Survey of the National Academy of sciences of Tajikistan (2009)
TM	Thai Seismic Monitoring Network (TM)
TR	Eastern Caribbean Seismograph Network
TT	Seismic Network of Tunisia
TW	Institute of Earth Sciences, Academia Sinica, Taiwan (1996)
TX	Bureau of Economic Geology, The University of Texas at Austin (2016)
UO	University of Oregon (1990)
US	Albuquerque Seismological Laboratory (ASL)/USGS (1990)
UU	University of Utah (1962)
UW	University of Washington (1963)
WA	West Central Argentina Network
WI	Institut De Physique Du Globe De Paris (IPGP) (2008)
WM	San Fernando Royal Naval Observatory (ROA) et al. (1996)
WU	The Southern Ontario Seismic Network (SOSN)
WY	University of Utah (1983)
XA	Paul Silver (1997), Kate Miller (2002)
XB	Douglas Wiens (1993)
XE	Douglas Christensen et al. (1999)
XF	Douglas Wiens (2012)
XI	Frank Vernon (1995)
XJ	Cynthia Ebinger (2013)
XR	Jim Ni et al. (1997)
XS	Stephane Rondenay (2006)
XW	Sylvie Leroy et al. (2009)

Table 2 continued: Seismological Networks used in CREW

XY	Susan Schwartz (1999),Steve Roecker and Ray Russo (2010)
XZ	Roger Hansen and Gary Pavlis (2005)
YC	Susan Beck et al. (2000),Anne Meltzer (2011)
YG	Carpathian Basins Project Regional Array (CBPRA)
YH	DANA (2012)
YI	Vadim Levin (2003)
YJ	Ethiopia-Afar Geoscientific Lithospheric Experiment (EAGLE)
YK	Coordinated Seismic Experiment in the Azores (COSEA)
YL	Anne Sheehan et al. (2001)
YO	Geoffrey A. Abers and Karen M. Fischer (2003)
YQ	Jim Gaherty et al. (2013)
YV	North East Atlantic Tomography (NEAT)
ZA	Michael West (2006)
ZC	Jay Pulliam (2013)
ZE	Cindy Ebinger (2007)
ZF	Afar Consortium Network (AFAR)
ZP	Andy Nyblade (2007)

REPORT
doi:10.26443/seismica.v3i1.1009

The SCEC/USGS Community Stress Drop Validation Study Using the 2019 Ridgecrest Earthquake Sequence

Annemarie Baltay *[1], **Rachel E. Abercrombie** [2], **Shanna Chu** [1], **Taka'aki Taira** [3]

[1]United States Geological Survey, Earthquake Science Center, Moffett Field, CA, USA, [2]Department of Earth & Environment, Boston University, Boston, MA, USA, [3]Berkeley Seismological Laboratory, UC Berkeley, Berkeley, CA, USA

Author contributions: *Conceptualization*: A. Baltay, R.E. Abercrombie. *Data Curation*: T. Taira. *Formal Analysis*: A. Baltay, R.E. Abercrombie, S. Chu. *Funding Acquisition*: A. Baltay, R.E. Abercrombie, T. Taira. *Project Administration*: A. Baltay, R.E. Abercrombie,. *Visualization*: A. Baltay, R.E. Abercrombie, S. Chu. *Writing – original draft*: A. Baltay, R. E. Abercrombie. *Writing – review & editing*: A. Baltay, R.E. Abercrombie, S. Chu, T. Taira.

Abstract We introduce a community stress drop validation study using the 2019 Ridgecrest, California, earthquake sequence, in which researchers are invited to use a common dataset to independently estimate comparable measurements using a variety of methods. Stress drop is the change in average shear stress on a fault during earthquake rupture, and as such is a key parameter in many ground motion, rupture simulation, and source physics problems in earthquake science. Spectral stress drop is commonly estimated by fitting the shape of the radiated energy spectrum, yet estimates for an individual earthquake made by different studies can vary hugely. In this community study, sponsored jointly by the U. S. Geological Survey and Southern/Statewide California Earthquake Center, we seek to understand the sources of variability and uncertainty in earthquake stress drop through quantitative comparison of submitted stress drops. The publicly available dataset consists of nearly 13,000 earthquakes of M1 to 7 from two weeks of the 2019 Ridgecrest sequence recorded on stations within 1-degree. As a community study, findings are shared through workshops and meetings and all are invited to join at any time, at any interest level.

Non-technical summary The stress release (or stress drop) during an earthquake provides information on how geologic forces are converted to radiated seismic energy when a fault ruptures, and the conditions under which an earthquake will continue to increase in size or trigger earthquakes nearby. Stress drop is also an important element of seismic hazard mapping and building design, since high stress drop earthquakes radiate more high frequency energy, resulting in stronger ground shaking. Unfortunately, stress drop estimates made in different studies have large systematic and random differences, implying that they are not as reliable as we need for use in ground motion prediction and earthquake source physics research. We introduce a Community Stress Drop Validation Study in which we invite all interested scientists from the international community to analyze the same earthquakes and compare and contrast their results. We use a public dataset of recordings of aftershocks of the 2019 Ridgecrest, California earthquake. Our aim is to understand where the differences and similarities in stress drop come from, and then work with the wider user community to develop improved methods for characterizing earthquake rupture and the resulting ground motions for more reliable and informed earthquake hazard forecasts.

Production Editor:
Gareth Funning
Handling Editor:
Stephen Hicks
Copy & Layout Editor:
Hannah F. Mark

Received:
June 23, 2023
Accepted:
April 25, 2024
Published:
May 22, 2024

1 Introduction

"What is earthquake stress drop, and what does it represent physically?" is a long-standing, open question in earthquake physics (e.g., Abercrombie, 2021). Seismologists and ground-motion modelers often mean dynamic stress drop, the change in shear stress driving earthquake faulting that goes into radiated seismic energy, which controls the amplitude and frequency content of ground shaking during earthquakes and is thus of great interest to ground-motion modelers and structural engineers. Geologists often mean static stress drop, the change in average stress resolved onto the fault before and after an earthquake rupture, which controls the mechanics of crustal deformation and should be related to slip on a fault, which can feed into earthquake occurrence statistics. In idealized, theoretical earthquake models, static and dynamic stress drops are equivalent: the dynamic high-frequency stress drop that can be measured from the radiated far-field seismogram is the same physical parameter as the static low-frequency stress drop that relates earthquake moment to rupture area.

To first order, this equivalency between various stress drop definitions and estimates has been observed, suggesting that earthquakes rupture in approximately the same way in a variety of geologic settings and over a wide range of magnitudes. This allows us to extrapolate current models and knowledge to predict ground-motion, slip, recurrence rates and other parameters to poorly recorded large-magnitude events, close distances, or new regions of interest. To improve our understanding of earthquake rupture dynamics, and de-

*Corresponding author: abaltay@usgs.gov

termine the factors that control earthquake rupture nucleation, propagation and arrest, we need to understand the real variation in earthquake stress drop (see Abercrombie, 2021).

Typically, seismologists estimate an average spectral stress drop $\Delta\sigma$ for an earthquake from the recorded Fourier frequency-amplitude spectrum by first fitting an ideal displacement source spectrum *u(f)*, with *f* the frequency, as a function of seismic moment (M_o) and corner frequency (f_c)

$$u(f) = \frac{M_0}{\left[1 + \left(\frac{f}{f_c}\right)^{n\gamma}\right]^{\frac{1}{\gamma}}} \tag{1}$$

where *n* is the high-frequency falloff rate and γ governs the shape near the corner. The commonly-used Brune (1970) model has *n*=2 and $\gamma = 1$. Then $\Delta\sigma$ is simply derived from the estimated corner frequency and moment, assuming a circular crack (Eshelby, 1957):

$$\Delta\sigma = cM_o\left(\frac{f_c}{k\beta}\right)^3 \tag{2}$$

where *c* is a constant accounting for rupture geometry (7/16 for a circular rupture) and k depends on the rupture velocity, wave type, and source model (typically 0.2-0.3, e.g., Brune, 1970; Madariaga, 1976; Sato and Hirasawa, 1973; Kaneko and Shearer, 2015). The corner frequency is inversely proportional to the wavelength of peak radiated energy from the source. Thus, stress drop can be thought of as the link between the low-frequency estimates of seismic moment and the high-frequency radiated energy assuming simple Brune-type circular crack models (Brune, 1970; Madariaga, 1976) in which corner frequency is inversely proportional to the rupture radius. Other source models are also possible, such as (Boatwright, 1978) with *n*=2 and $\gamma = 2$ in Eq. 1) or double corner models where the low-frequency f_c is related to the source duration and hence dynamic stress drop discussed herein, and the higher f_c is related to a secondary process such as rise time, starting or stopping phases, or a dynamic weakening process (e.g., Denolle and Shearer, 2016),

Throughout our study, and in this paper, we focus on this widely used spectral estimate of stress drop, whether it comes directly from the corner frequency or a related parameter, such as duration, energy, or high-frequency ground motion. While the alternate name of "stress parameter" is in use to describe the source spectral shape in ground motion modeling (i.e., Atkinson and Beresnev, 1997), due to the large uncertainties and difficulties relating it to any actual stress drop in the earth, here we use the simple term "spectral stress drop" for spectral estimates to match current practice. Apparent stress, defined as rigidity times the ratio of broadband radiated seismic energy to moment, is theoretically a more model-independent estimate of the stress drop (e.g., Ide and Beroza, 2001; Baltay et al., 2010). In practice, however, accurate measurement of radiated broadband energy is challenging as it requires extrapolation to high frequencies, and often depends on the same spectral modeling as the spectral stress drop estimates, because of the need to model high-frequency attenuation and other path effects. These measurements are especially difficult at the higher frequencies required to quantify radiated energy of smaller earthquakes (e.g. Abercrombie, 1995; Ide and Beroza, 2001; Abercrombie, 2021). This spectral stress drop is an average stress drop over an earthquake rupture, and the relationship between that average and time- and space-varying stress drop on a fault is not always well resolved (Noda et al., 2013). Similarly, the details of the relationship between this seismological spectral stress drop and the actual stress release on a fault or numerical simulations are poorly understood (e.g., Kaneko and Shearer, 2015; Ji et al., 2022). Before we can attempt to connect all these parameters, we need to first ensure our estimate of the spectral stress drop is reliable and reproducible; this is the aim of the community study.

The ease with which it can be measured and its importance for both earthquake physics and high-frequency ground-motion modeling, have led to spectral stress drop becoming a frequent subject of study worldwide (e.g., Aki, 1967; Hanks, 1977; Abercrombie, 1995; Ide and Beroza, 2001; Baltay et al., 2011; Abercrombie et al., 2016). However, for as long as stress drop has been measured, it has been a topic of debate, as stress drop estimates are rife with uncertainties and appear highly variable (Cotton et al., 2013; Abercrombie, 2021).

While we often observe an approximately constant range of stress drop over a wide range of earthquake magnitudes, the variation within individual studies can be three orders of magnitude (e.g., Figure 1). How much of this is due to measurement uncertainty, and how much to real inter-event variation is unknown. For individual earthquakes, stress drops estimated by different researchers or using different methods rarely agree (e.g., Abercrombie, 2013; Pennington et al., 2021), with differences between estimates larger than the reported uncertainties, implying that calculated uncertainties of at least some approaches must be significantly underestimated. On a larger scale, it is still an open question as to whether stress drop scales with magnitude (e.g., Baltay et al., 2010; Bindi et al., 2020), depth (e.g., Hardebeck and Aron, 2009; Trugman and Shearer, 2017; Abercrombie et al., 2021), faulting regime or tectonic setting (e.g., Allmann and Shearer, 2009; Boyd et al., 2017; Huang et al., 2017), or even nature and extent of dynamic weakening or thermal pressurization (Beeler et al., 2012; Nielsen et al., 2016; Rice, 2006). However, the large scatter currently obscures these trends, so for stress drop to be most reliably used both to understand rupture physics and in models and simulations, we need to understand how physical processes, methodological differences, and data processing artifacts contribute to these variations.

Various studies have investigated the effects of differences in methods or data selection, including Shearer et al. (2019), Goertz-Allmann and Edwards (2013), Abercrombie (2015), Chen and Abercrombie (2020), Pennington et al. (2021), and Shible et al. (2022); Abercrombie (2021) provides a broad review of the difficulties, uncertainties and methods in stress drop estimation and comparison. These studies found that although methodological differences can lead to some systematic biases, the main differences come from the simplifying assumptions and model parameterization, and the limited quality and quantity of the data. Objectively determining the most reliable approaches for calculating stress drop, and more representative estimates of uncertainties, is beyond the abilities of any individual group.

Awareness of the need for a community-wide study to resolve these discrepancies has been growing over the years (e.g., Baltay et al., 2017; E.C.G.S. Workshop, 2012). Therefore, this Community Stress Drop Validation Study was initiated by co-leads Annemarie Baltay and Rachel Abercrombie in 2021, with support from the U.S. Geological Survey (USGS) and Southern California Earthquake Center (SCEC); the Statewide California Earthquake Center (SCEC) continues to support this project in 2024. The goals of this group are to understand: (1) the sources of agreement or difference between different methods and data sets used in estimating stress drop, (2) how physical attributes of the earthquake source affect the variability or degree of agreement of those estimates, and (3) ultimately, what is the best path forward for measuring stress drop and characterizing the high frequency radiation for various end-user needs. The 2019 Ridgecrest earthquake sequence provides the perfect dataset for such a comparative study.

2 Research priorities and organization

2.1 Research priorities

The goals of this Community Stress Drop Validation Study are to understand the nature and causes of variability and uncertainty in spectral earthquake stress drop estimates and how physical effects, random errors, differing data sets and methodological variability may contribute to these discrepancies, so that we best understand and account for these uncertainties.

Our specific research priorities are to:

1. Understand how different methods and assumptions lead to variations in estimated stress drop and predicted high frequency radiation. Do certain methods highlight different frequency aspects of the source? How do data selection and preprocessing affect the results? How are different analysts implementing methods?

2. Determine how variations in the estimated spectral stress drops reflect physical variations in earthquake source processes or material properties. Do simpler or smoother events yield more agreement between stress drop estimates while complex events show more variability? How do these stress drop estimates depend on the physical size, depth, location or tectonic setting of the earthquake?

3. Develop best practices for estimating a measure of spectral stress drop that can reliably be used in ground motion and hazard modeling, and by the wide community seeking to understand earthquake source physics and dynamic rupture processes (including laboratory work and numerical modeling). Ultimately, the best way to estimate stress drop may vary between events depending on factors such as its tectonic setting, inferred rheological properties and rupture behavior, but can we develop a baseline method that is consistent for a particular type of earthquake?

2.2 Study organization

The overall process for the Stress Drop Validation Study is to: provide and distribute a common dataset from the 2019 Ridgecrest sequence; solicit community researchers to carry out analyses of stress drop, or related parameters, for those events; return results of these analyses to the project leads for systematic comparison and meta-analysis; and discuss and disseminate these findings through scientific conferences, workshop discussions, and publications. In addition to attracting participants to make measurements, a major aim of the group is to engage end users to promote informed use of observational measurements with understanding of the uncertainties, and also assist in developing and making the most useful measurements needed to advance hazards and earthquake physics research. This study is envisioned as an iterative and community-driven process to help the seismological community strengthen our understanding of stress drop variability and uncertainty, and what it can tell us about the physics of earthquake rupture and the resulting ground motions.

This project has focused on building community and encouraging collaboration between participants to stimulate validation efforts, leading to sub-groups performing comparative analysis, and investigating the effects of method variations (e.g., Bindi et al., 2023a,b). Through support from SCEC, we have hosted three virtual workshops in November 2021, January 2023 and January 2024, and one in-person workshop at the SCEC Annual Meeting in September 2022. The virtual workshops have attracted over 100 participants each from 20 countries and all continents (except Antarctica), while the more focused in-person workshop was 30 participants. At each workshop, recent results and meta-analysis are shared, and the group discusses future directions including also hearing from stress drop users, rather than just analysts. At the most recent January 2024 workshop, for example, we discussed creation and analysis of synthetic datasets, hearing about several different methods for simulating waveforms (full workshop reports can be found at https://www.scec.org/research/stress-drop-validation). In between workshops, we hold ~monthly video-conference calls for community building and validation activities, which are typically held at two different times in the same day to encourage and enable global contributions; we currently have broad geographical participation.

3 Current validation study: 2019 Ridgecrest earthquake sequence

The current community stress drop validation study is focused on the 2019 Ridgecrest earthquake sequence

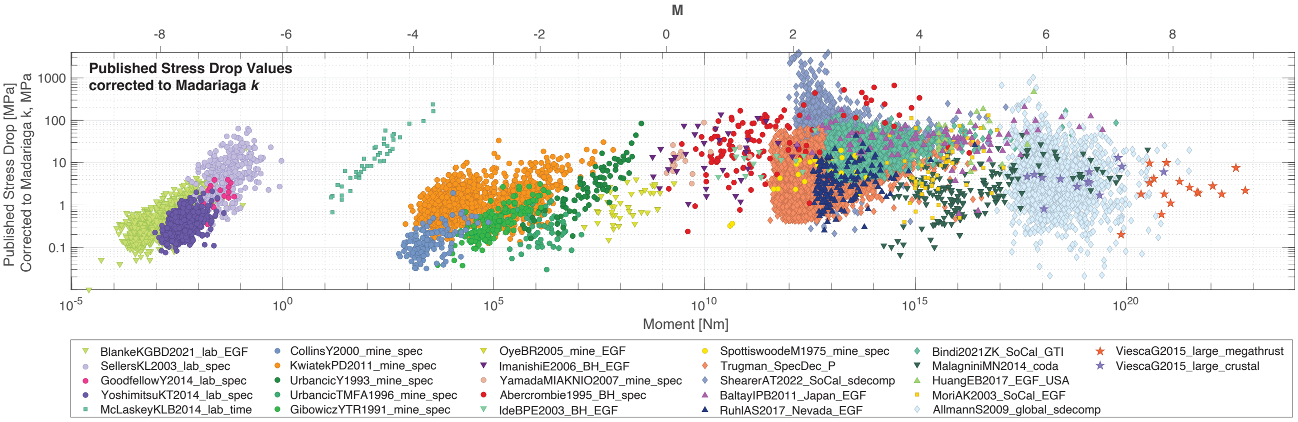


Figure 1 Published stress drop compilation showing stress drops versus magnitude for global earthquakes across a wide range of magnitudes. For each study, the stress drop is corrected assuming the k=0.21 value from Madariaga (1976), to avoid discrepancies purely from author choice of k. While there is very large scatter between and across studies, stress drops are generally bounded between 0.1 and 100 MPa, for events ranging from acoustic emissions recorded in the lab and during mine-break experiments, (Yoshimitsu et al., 2014; Sellers et al., 2003; Kwiatek et al., 2011; Spottiswoode and McGarr, 1975; Urbancic et al., 1996; Urbancic and Young, 1993; Gibowicz et al., 1991; Collins and Young, 2000; Oye et al., 2005; Yamada et al., 2007; Goodfellow and Young, 2014; Blanke et al., 2021; McLaskey et al., 2014), to regional studies (Abercrombie, 1995; Imanishi and Ellsworth, 2006; Trugman, 2020; Shearer et al., 2022; Bindi et al., 2021; Malagnini et al., 2013; Baltay et al., 2011; Huang et al., 2017; Ide et al., 2003; Mori et al., 2003; Baltay et al., 2010; Ruhl et al., 2017) and global compilations (Allmann and Shearer, 2009; Viesca and Garagash, 2015).

using a set of common waveforms. The study is divided into two main research activities: 1) Independent analysis of stress drop for the Ridgecrest sequence by researchers, and submission to the group validation repository; and 2) Meta-analysis to compare the submitted results. The study is inclusive and iterative, in that any researchers may join at any time to provide their estimates of stress drops; then as a group we compare all stress drop estimates and refine the stated problem and narrow the data set to best achieve our goals. Individual researchers will then repeat some aspects of their analysis with newfound insight and using a more limited data set.

We have created and provide a common data set for this study, including waveforms and metadata, available for download through the Southern California Earthquake Data Center: https://scedc.caltech.edu/data/stressdrop-ridgecrest.html, where a "Quick-reference guide" is also posted for more information on the waveform data. This dataset consists of ~13,000 earthquakes of magnitude 1+ over two weeks from July 4 until July 17 (Figure 2). This contains the M7.1 and M6.4 Ridgecrest mainshocks, three M5 earthquakes and 86 M4 events. This two-week window was chosen to avoid introducing selection biases yet retain a set of earthquakes sufficient for the wide variety of expected stress drop analyses. It is unlikely that any individual contributor will analyze all the earthquakes, and the approaches of different groups will be suitable for different subsets. To increase comparison, we have selected a subset of 55 events, by choosing well recorded events over a range of magnitudes from 2 to 4.5, at a range of depths and along different parts of the rupture. We ask researchers to prioritize these events in their analysis, if possible.

3.1 Waveform data

The provided data are recorded on 107 local and regional stations within 1-degree (~110km) of each epicenter, and consist of broadband velocimeter, accelerometer and geophone instruments—including both horizontal and vertical components. Data come from the Southern California Data Center (SCEDC), International Research Institutions for Seismology Data Management Center, and the Northern California Earthquake Data Center. We included network codes CE, CI, GS, NN, NP, PB, SN, and ZY but excluded the nodal network 3J, and used channels HH (up to 200 sps) and CH (> than 200sps) for broadband, HN (<200 sps) and CN (>200 sps) for accelerometers, and EP (<200 sps), EH (S200sps) and DP (>200 sps) for geophones (see Data and Code Availability section); in each case the channel with the highest sampling rate is chosen for co-located instruments. The length of each record is proportional to the magnitude, with the record starting 15s before the origin time (OT) and ending 60s after for M1; for the M6+ the records start 90 before OT and end 310s after. The waveforms are provided in miniSEED format and can be directly downloaded as tar files grouped by magnitude, to reduce file size for any one archive. Within each tar file is a folder for each earthquake; within that folder is a list of stations for that event, accompanying response information (SAC pole-zero files) and StationXML metadata. The ObsPy (Beyreuther et al., 2010) script used to create this dataset is available for use as well, either to facilitate direct download of the waveforms, or to adjust any of the parameters.

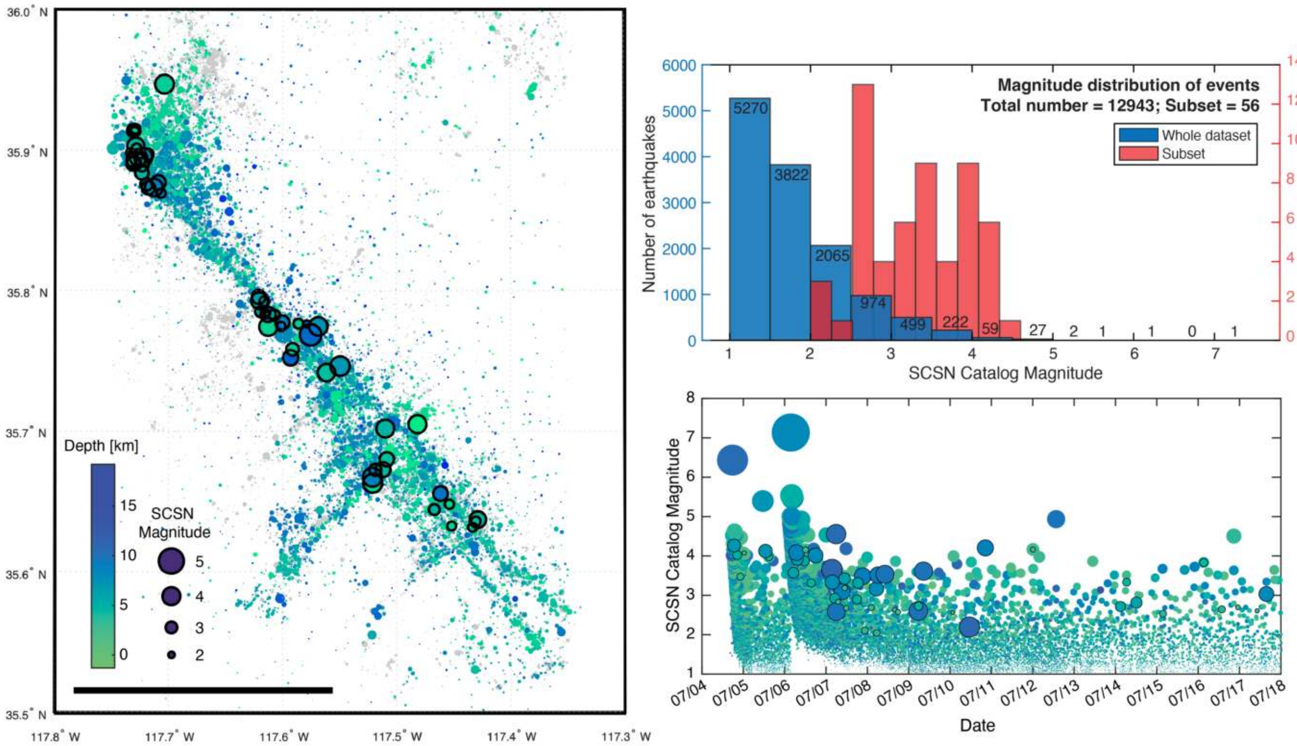


Figure 2 Event locations (left), magnitude (top right) and time-vs.-magnitude distribution (bottom right). Inset map shows location of Ridgecrest region (red box) within the state of California. Entire two-week relocated event catalog of ~13,000 earthquakes by Trugman (2020), shown in circles colored by depth and sized by magnitude and in blue histogram bars. Subset of 55 events for focused study shown encircled in black, and in red histogram bars.

3.2 Metadata

Along with the waveform data, we provide several metadata to assist in analysis, and remove unnecessary sources of variation between results.

- *Earthquake Catalog.* Full earthquake catalog with SCSN magnitudes and relocations from Trugman (2020).
- *P- and S-wave phase picks.* Initial P- and S-wave phase picks for each record (although if a method requires improved picks, participants are free to adjust or repick the data) through two methods: The first are the SCEDC phase picks, which are not available for all events or all stations; the second are theoretical travel time calculations using a 1D velocity model. Both sets of phase-picks are included batched into the .tar files with the waveforms.
- *V_s30 station estimates.* Time-averaged shear-wave velocity in the upper 30m (V_s30) for each station. The V_s30 values are preferentially measured, as reported by Yong et al. (2013); if direct measurements are not available then V_s30 is estimated based on the mosaic proxy of Heath et al. (2020).
- *Ridgecrest 1D velocity model.* A simple 1D velocity model for those wanting depth-dependent rupture velocity correction, developed by White (2021), by combining and discretizing the models from Lin et al. (2007) (25% weight), Zhang and Lin (2014) (25% weight) and White et al. (2021) (50% weight).

4 Earthquake stress drop analysis

4.1 Individual stress drop analysis

Throughout the study, we solicit submissions of stress drop or other source parameter estimates (source duration, finite fault inversions, high-frequency energy, etc.) in a defined spreadsheet format from the community via the email distribution list. New and updated submissions of results and participation are still encouraged, especially from students, early career, and international (non-US) participants. To participate in the community study, we ask that participants be willing to provide their analyses potentially ahead of publication, so that they can iterate on methods and analysis. This allows them to understand and isolate sources of discrepancy or variability in their analyses, which will both improve the quality and impact of their own publications and eventually better inform other community members about alternative approaches and possible outcomes. Submission of the results is made only to the authors (study PIs), to ensure confidentiality of the results. Participants are asked for their permission before any results are shown to the larger group or included in presentations. To date, we have received 47 unique submissions from 20 research groups.

The common methods of estimating spectral stress drop, and their limitations, are reviewed by Abercrombie (2021). The original, simplest method of fitting individual earthquake spectra to determine source, path and site (e.g., Thatcher and Hanks, 1973) is still in use (e.g., Kemna et al., 2021) but has proven to be poorly constrained (e.g., Ko et al., 2012). When sufficient quantity and quality of recordings are available, variations on two distinct approaches are currently preferred to isolate the source, and estimate corner frequency, source duration or stress drop, and they both can use body or coda waves (see Abercrombie, 2021). Variations and combinations of these have been used by participants in the Community Study to date, and the authors cited below have all submitted preliminary results at the time of writing.

1. Spectral Decomposition / Generalized Inversion: A range of different inversion strategies are now in use, commonly known as spectral decomposition or generalized inversion techniques (GIT), for example, Shearer et al. (2006), Chen and Shearer (2011), Pennington et al. (2021), Trugman (2020), Bindi et al. (2021), Devin et al. (2021), Vandevert et al. (2022). These inversions simultaneously invert large numbers of earthquakes and stations for stability to obtain single, station-averaged values. Obtaining absolute values of source parameters, including earthquake magnitude, requires assumption of a source model (typically a Brune-type spectrum) or a constraint on the average site effect, for example, assuming a flat response at a reference rock site. These inversions also incorporate an azimuthally independent attenuation structure, which is assumed to be either homogeneous (constant) or a simple function of travel time.

2. Empirical Green's Function (EGF) Analysis: In this empirical approach, a small, co-located earthquake is used as an EGF to remove path and site effects from the spectrum or seismogram of a larger target earthquake. The deconvolution requires no assumptions about path or site effects, and can be applied to individual pairs of events, at individual stations to enable investigation of azimuthal variation in the source radiation and path effects. It requires an independent estimate of seismic moment of one or both events, a source model with which to fit the corner frequencies (could be one as given in Eq. 1 or an assumption that the EGF event is flat to displacement in the relevant frequency range), and depends on the availability of an appropriate, well-recorded EGF earthquake, which significantly limits the number of events that can be studied using this method. The results also depend on the correctness of the EGF assumption, and research into the effects of EGF choice is ongoing (e.g., Abercrombie et al., 2016). Spectral ratios are usually calculated by direct division of the amplitude spectra, but the source time functions can be calculated either by complex spectral division or by time-domain inversion. To obtain source parameters, the spectral ratios are fit with a simple Brune-source model (e.g., Abercrombie et al., 2020; Kemna et al., 2021; Liu et al., 2020; Ruhl et al., 2017; Boyd et al., 2017; Chen and Shearer, 2011; Mayeda et al., 2007). Alternatively, a finite fault or other inversion can be used to model the source time functions (e.g., Dreger et al., 2021; Fan et al., 2022).

Many approaches in common usage are variations and combinations of these two. For example, the coda calibration tool approach (Mayeda et al., 2003) uses coda spectral ratios of one or two calibration events to constrain the path and site corrections for other individual events and Eulenfeld et al. (2021) combine coda wave estimation of attenuation with a generalised inversion of the direct wave spectra. Kemna et al. (2021) and Boyd et al. (2017) use cluster-based approaches to constrain individual spectral fitting and spectral ratio modelling, respectively. Supino et al. (2019) develop a probabilistic framework for the inversion, and Satriano (2022) uses an iterative approach, first fitting individual body wave spectra then refining the fits with station-specific average constraints.

Several methods are distinct from the two main approaches, such as Knudson et al. (2023) and Al-Ismail et al. (2023), who calculate the amplitude spectra at individual points from the amplitudes of narrow-band filtered seismograms. Baltay et al. (2019) use ground-motion intensities to directly estimate stress drop, and Ji et al. (2022) estimate stress drop based on radiated energy.

4.2 Initial results and meta analysis

Direct comparison of the stress drops submitted to the Community Stress Drop Validation study so far reveals considerable scatter, but some stronger correlation between results using similar methods. The relative variations between different earthquakes are more consistent across the various studies, than are the absolute values, in line with the results of Pennington et al. (2021). We also observe some systematic magnitude- and depth-dependent overall offsets between different authors' submissions. Overall, we observe a stronger increase of stress drop with earthquake source depth for methods that do not allow travel-time dependent attenuation to vary with source depth. This implies that some of the increased stress drop with depth may be due to tradeoffs with attenuation and near-source structure, consistent with the results of Abercrombie et al. (2021).

To date, we have focused primarily on the estimates of corner frequency, and many methods also estimate seismic moment. We see large scatter in estimated corner frequency and also some considerable scatter in moment; some studies find an increase in spectral stress drop with increasing moment, but a constant stress drop is within the uncertainties for most, if not all, results. Whether any magnitude dependence to stress drop is real, or a consequence of the frequency bandwidth, simplistic assumptions and method selections used (e.g., Abercrombie, 2021) is not yet clear.

Of the 47 unique stress drop submissions received so far, 21 are published (Figure 3): Trugman (2020), Shearer et al. (2022), Bindi et al. (2021) and Bindi et al. (2023b), the latter of which included 18 variations using different parameters. These results all show relatively

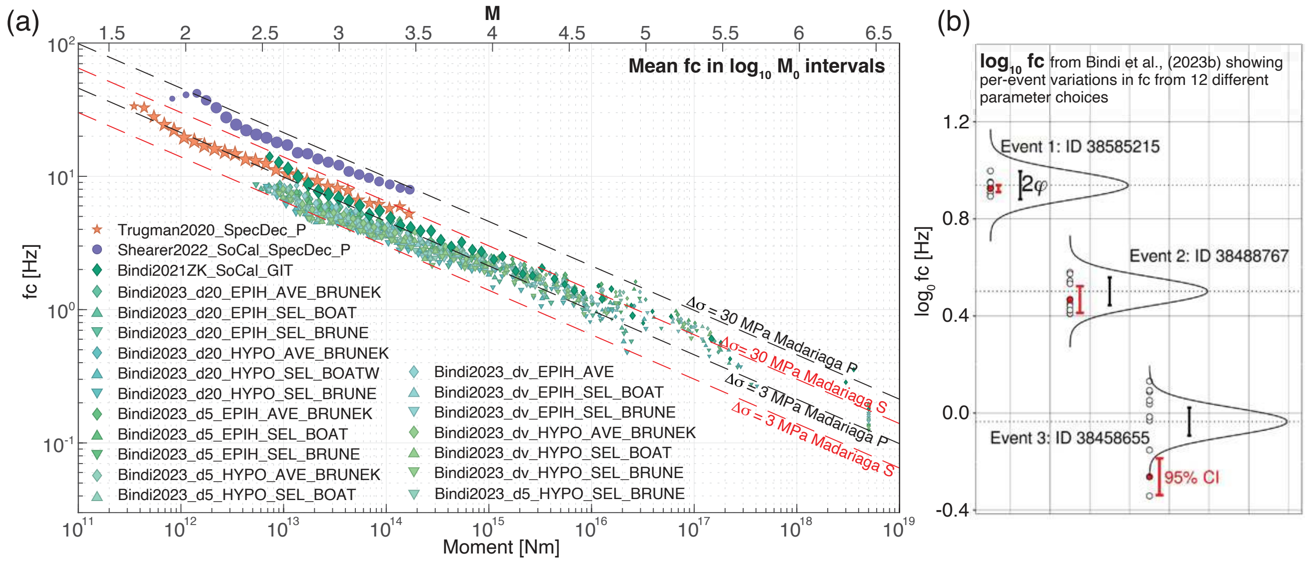


Figure 3 Comparison of published corner frequency results as part of the Community Stress Drop Validation Study. (a) Estimated corner frequency vs. estimated moment from Trugman (2020), Shearer et al. (2022), Bindi et al. (2021) and Bindi et al. (2023b), with dashed diagonal lines showing constant values of stress drop under the assumption of a Madariaga (1976) *k=0.21* for both P and S waves. (b) Comparison of resultant corner frequency from the 12 different parameter choices using a Brune (1970) spectra, from Bindi et al. (2023b) for three representative events (Event 1 M2.7; Event 2 M3.3; Event 3 M4.2). For the case shown in red filled circle and bar, the 95% confidence interval on that estimate sometimes doesn't overlap with the other estimates given other parameter choices. Figure (b) reproduced from Bindi et al. (2023b) Figure 6b.

constant stress drop scaling with magnitude (i.e., falling along a line of constant stress drop) and are recovering stress drops in a range of 3 to 30 Mpa, upholding expectations for regions in California. All these published results are large scale spectral decomposition/-generalized inversion technique methods on the Ridgecrest 2019 sequence, so although these methods are very similar, there are significant systematic differences between them. Corner frequencies derived from P waves should be larger than those from S waves. While we see that estimates from both Trugman (2020) and Shearer et al. (2022), who use P waves, are indeed larger than those from the Bindi et al. (2021, 2023a,b) studies which all use S waves, the difference is larger than predicted by theoretical models; there is still significant offset between the two P-wave studies, similar to the range in the S-wave estimates obtained using different method variations. We need to further understand if there is a physical or simply methodological reason behind these discrepancies, and comparative studies such as Bindi et al. (2023a,b) are extremely valuable in determining the real systematic and random errors.

Bindi et al. (2023a,b) iterated over several parameter choices, including spectral window duration, source depth dependent or independent attenuation, different approaches for normalizing the site constraint, and fitting with a Brune (1970) or Boatwright (1978) spectral shape. For some specific events, the different corner frequencies estimated over these different iterations show good agreement (i.e., Event 1 in Figure 3b) while some events show large disagreement (i.e., Event 3). When considering the standard error of 95% confidence on one iteration, shown as the red bar in Figure 3b, sometimes the standard error encompasses the variability of the various iterations and sometimes does not, implying that method choices and assumptions can lead to wider variation than the formal errors in a single preferred approach, that are typically published. It remains to be seen if there are physical predictors or complexity that might indicate when estimated corner frequencies will agree or not.

We also find that a major source of disagreement stems from estimated seismic moments submitted for the same events. Many methods that generate displacement source spectra fit an estimated moment as well as a corner frequency, typically using a Brune (1970) spectra and fitting for the seismic moment M_0 as well as the corner frequency (Equation 1). Thus, there is inherent tradeoff in the two fitted parameters M_0 and f_c and we observe almost as much variability in submitted moments, as do submitted corner frequencies. We also convert the submitted moments to moment magnitude as $\mathbf{M} = \frac{2}{3}(log_{10}M_0 - 9.05)$, following Hanks and Kanamori (1979), and find both scatter and systematic differences between these $\mathbf{M}$ and the catalog moment magnitudes. The relationship between catalog measurements of local magnitude, coda magnitude, etc. and moment magnitude below M~4 is not simple (e.g., Hanks and Boore, 1984), and an incomplete understanding of magnitude can cause systematic bias in source parameter estimates as well as in statistical estimation of b-value, for example. However, the results compiled in this study provide a unique opportunity to improve moment-magnitude relationships in Southern California, and also potentially lead to a more physics based revised local magnitude scale (Mlr, https://scedc.caltech.edu/eq-catalogs/change-history.html).

5 Outlook

From the initial submitted and published results, it is apparent that more detailed analysis will improve understanding of why different methods and assumptions for estimating stress drop, or different researchers applying similar methods, yield different results. There are many places where workflows can differ, and so isolating how different choices affect the estimates, and which have the largest effects may improve coherency of results. Toward this end, it is encouraging to see many researchers within our community starting to study the sensitivity of estimated parameters to the various input choices (e.g., Bindi et al., 2023a), and initiating collaborations to compare approaches (e.g., Morasca et al., 2022).

To isolate and quantify specific sources of variability, we are conducting benchmark studies. In the first study, we are testing how results from different researchers vary even when they start out with the same source spectra. We have found that the variability in the benchmark fitting with fixed source spectra is about 3-10 times smaller when compared to overall results, indicating that spectral fitting is a small but relevant portion of the overall variability. Future benchmarks will enable us to isolate the effects of window length and frequency band selection, and other pre-processing choices. Providing an augmented dataset to include a processed ground-motion style flat file will facilitate participation of ground motion researchers in the study.

Joining the ongoing Community Stress Drop Validation Study is straightforward: one can download the data and perform analysis for stress drop, corner frequency or other source parameters, become involved in the meta-analysis to compare different results, or simply join in workshops to learn more about stress drop analysis or understand better how seismological measurements can constrain or inform their own research (https://www.scec.org/research/stress-drop-validation, or contact the authors). Even after this stage of the study is completed and published, the data and study description will enable future researchers to test and compare new methods and codes to existing methods.

Acknowledgements

The authors are very thankful for the community support in starting this effort, especially all those researchers who have contributed stress drop estimates, those who have participated in workshops and meetings, and everyone who has been a sounding board for this work, especially Bill Ellsworth, Tom Hanks, Elizabeth Cochran, Colin Pennington, and Greg Beroza. The authors thank the Southern California Earthquake Center and the new Statewide California Earthquake Center (SCEC) for providing the Technical Activity Group framework which made this effort a fully supported project; to the SCEC meetings team, especially Tran Huynh and Edric Pauk, without whom none of our workshops would be possible; the Southern California Earthquake Data Center (SCECDC) and others who helped with making the data set and meta-data and the whole SCEC community for their support. We thank Steve Hickman, German Prieto and editor Stephen Hicks for their thorough reviews of this manuscript. R.E. Abercrombie thanks the U.S. Geological Survey for funding to support this effort. S. Chu thanks Pacific Gas and Electric for support for this work. This research was supported by SCEC (awards 21083, 21114, 22101, 22042, 23107 and 23108 including salary support for T. Taira and financial support for the 2022 September workshop). SCEC is funded by NSF Cooperative Agreement EAR-1600087 & USGS Cooperative Agreement G17AC00047. SCEC Contribution #13444. Any use of trade, firm, or product names does not imply endorsement by the U.S. Government.

Data and code availability

All of the waveform data and metadata referenced herein are publicly available and accessed from the Southern California Earthquake Data Center (SCEDC), Incorporated Research Institutes for Seismology (IRIS) and Northern California Earthquake Data Center (NCEDC; doi:10.7932/NCEDC). The SCEDC (http://scec.org) is funded by National Science Foundation (NSF) Cooperative Agreement EAR-1600087 and USGS Cooperative Agreement G17AC00047. IRIS Data Services are funded through the Seismological Facilities for the Advancement of Geoscience and EarthScope (SAGE) Proposal of the NSF under Cooperative Agreement EAR-1261681. Networks that provided data are: CE (California Geological Survey, 1972); CI (California Institute of Technology and United States Geological Survey Pasadena, 1926); GS (Albuquerque Seismological Laboratory, 1980); NN (University of Nevada, Reno, 1971); NP (U.S. Geological Survey, 1931); PB (https://www.fdsn.org/networks/detail/PB/); SN (University of Nevada, Reno, 1992); and ZY https://www.fdsn.org/networks/detail/ZY_1990/(Cochran et al., 2020). All data referenced here are available at the Southern California Earthquake Data Center (SCEDC, 2013) at https://scedc.caltech.edu/data/stressdrop-ridgecrest.html, where a "Quick-reference guide" is also posted for more information on the waveform data.

Competing interests

No competing interests.

References

Abercrombie, R. E. Earthquake source scaling relationships from −1 to 5 ML using seismograms recorded at 2.5-km depth. *Journal of Geophysical Research: Solid Earth*, 100(B12):24015–24036, Dec. 1995. doi: 10.1029/95jb02397.

Abercrombie, R. E. Comparison of direct and coda wave stress drop measurements for the Wells, Nevada, earthquake sequence. *Journal of Geophysical Research: Solid Earth*, 118(4):1458–1470, Apr. 2013. doi: 10.1029/2012jb009638.

Abercrombie, R. E. Investigating uncertainties in empirical Green's function analysis of earthquake source parameters. *Journal of Geophysical Research: Solid Earth*, 120(6):4263–4277, June 2015. doi: 10.1002/2015jb011984.

Abercrombie, R. E. Resolution and uncertainties in estimates of earthquake stress drop and energy release. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2196):20200131, Mar. 2021. doi: 10.1098/rsta.2020.0131.

Abercrombie, R. E., Bannister, S., Ristau, J., and Doser, D. Variability of earthquake stress drop in a subduction setting, the Hikurangi Margin, New Zealand. *Geophysical Journal International*, 208(1):306–320, Oct. 2016. doi: 10.1093/gji/ggw393.

Abercrombie, R. E., Chen, X., and Zhang, J. Repeating Earthquakes With Remarkably Repeatable Ruptures on the San Andreas Fault at Parkfield. *Geophysical Research Letters*, 47(23), Dec. 2020. doi: 10.1029/2020gl089820.

Abercrombie, R. E., Trugman, D. T., Shearer, P. M., Chen, X., Zhang, J., Pennington, C. N., Hardebeck, J. L., Goebel, T. H. W., and Ruhl, C. J. Does Earthquake Stress Drop Increase With Depth in the Crust? *Journal of Geophysical Research: Solid Earth*, 126(10), Oct. 2021. doi: 10.1029/2021jb022314.

Aki, K. Scaling law of seismic spectrum. *Journal of Geophysical Research*, 72(4):1217–1231, Feb. 1967. doi: 10.1029/jz072i004p01217.

Al-Ismail, F., Ellsworth, W. L., and Beroza, G. C. A Time-Domain Approach for Accurate Spectral Source Estimation with Application to Ridgecrest, California, Earthquakes. *Bulletin of the Seismological Society of America*, 113(3):1091–1101, Mar. 2023. doi: 10.1785/0120220228.

Albuquerque Seismological Laboratory. GS, International Federation of Digital Seismograph Networks, 1980. doi: 10.7914/SN/GS.

Allmann, B. P. and Shearer, P. M. Global variations of stress drop for moderate to large earthquakes. *Journal of Geophysical Research: Solid Earth*, 114(B1), Jan. 2009. doi: 10.1029/2008jb005821.

Atkinson, G. M. and Beresnev, I. Don't Call it Stress Drop. *Seismological Research Letters*, 68(1):3–4, Jan. 1997. doi: 10.1785/gssrl.68.1.3.

Baltay, A., Prieto, G., and Beroza, G. C. Radiated seismic energy from coda measurements and no scaling in apparent stress with seismic moment. *Journal of Geophysical Research: Solid Earth*, 115(B8), Aug. 2010. doi: 10.1029/2009jb006736.

Baltay, A., Ide, S., Prieto, G., and Beroza, G. Variability in earthquake stress drop and apparent stress. *Geophysical Research Letters*, 38(6), Mar. 2011. doi: 10.1029/2011gl046698.

Baltay, A., Ellsworth, W., Schoenball, M., and Beroza, G. Proposed Community Stress Drop Validation Experiment. In *SCEC Annual Meeting*, Palm Springs, CA, 2017.

Baltay, A. S., Hanks, T. C., and Abrahamson, N. A. Earthquake Stress Drop and Arias Intensity. *Journal of Geophysical Research: Solid Earth*, 124(4):3838–3852, Apr. 2019. doi: 10.1029/2018jb016753.

Beeler, N., Kilgore, B., McGarr, A., Fletcher, J., Evans, J., and Baker, S. R. Observed source parameters for dynamic rupture with non-uniform initial stress and relatively high fracture energy. *Journal of Structural Geology*, 38:77–89, May 2012. doi: 10.1016/j.jsg.2011.11.013.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., and Wassermann, J. ObsPy: A Python Toolbox for Seismology. *Seismological Research Letters*, 81(3):530–533, May 2010. doi: 10.1785/gssrl.81.3.530.

Bindi, D., Spallarossa, D., Picozzi, M., and Morasca, P. Reliability of Source Parameters for Small Events in Central Italy: Insights from Spectral Decomposition Analysis Applied to Both Synthetic and Real Data. *Bulletin of the Seismological Society of America*, 110(6):3139–3157, July 2020. doi: 10.1785/0120200126.

Bindi, D., Zaccarelli, R., and Kotha, S. R. Local and Moment Magnitude Analysis in the Ridgecrest Region, California: Impact on Interevent Ground-Motion Variability. *Bulletin of the Seismological Society of America*, 111(1), 2021. doi: 10.1785/0120200227.

Bindi, D., Spallarossa, D., Picozzi, M., Oth, A., Morasca, P., and Mayeda, K. The Community Stress-Drop Validation Study—Part II: Uncertainties of the Source Parameters and Stress Drop Analysis. *Seismological Research Letters*, May 2023a. doi: 10.1785/0220230020.

Bindi, D., Spallarossa, D., Picozzi, M., Oth, A., Morasca, P., and Mayeda, K. The Community Stress-Drop Validation Study—Part I: Source, Propagation, and Site Decomposition of Fourier Spectra. *Seismological Research Letters*, 94(4), 2023b. doi: 10.1785/0220230019.

Blanke, A., Kwiatek, G., Goebel, T., Bohnhoff, M., and Dresen, G. Stress drop–magnitude dependence of acoustic emissions during laboratory stick-slip. *Geophysical Journal International*, 224 (2), 2021. doi: 10.1093/gji/ggaa524.

Boatwright, J. Detailed Spectra Analysis of Two Small New York State Earthquakes. *Bulletin of the Seismological Society of America*, 68(4), 1978. doi: 10.1785/BSSA0680041117.

Boyd, O. S., McNamara, D. E., Hartzell, S., and Choy, G. Influence of Lithostatic Stress on Earthquake Stress Drops in North America. *Bulletin of the Seismological Society of America*, 107(2):856–868, Feb. 2017. doi: 10.1785/0120160219.

Brune, J. N. Tectonic stress and the spectra of seismic shear waves from earthquakes. *Journal of Geophysical Research*, 75(26): 4997–5009, Sept. 1970. doi: 10.1029/jb075i026p04997.

California Geological Survey. California Strong Motion Instrumentation Program, 1972. doi: 10.7914/B34Q-BB70.

California Institute of Technology and United States Geological Survey Pasadena. CI, International Federation of Digital Seismograph Networks, 1926. doi: 10.7914/SN/CI.

Chen, X. and Abercrombie, R. E. Improved approach for stress drop estimation and its application to an induced earthquake sequence in Oklahoma. *Geophysical Journal International*, 223 (1):233–253, June 2020. doi: 10.1093/gji/ggaa316.

Chen, X. and Shearer, P. M. Comprehensive analysis of earthquake source spectra and swarms in the Salton Trough, California. *Journal of Geophysical Research*, 116(B9), Sept. 2011. doi: 10.1029/2011jb008263.

Cochran, E. S., Wolin, E., McNamara, D. E., Yong, A., Wilson, D., Alvarez, M., van der Elst, N., McClain, A., and Steidl, J. The U.S. Geological Survey's Rapid Seismic Array Deployment for the 2019 Ridgecrest Earthquake Sequence. *Seismological Research Letters*, 91(4):1952–1960, Jan. 2020. doi: 10.1785/0220190296.

Collins, D. and Young, R. Lithological Controls on Seismicity in Granitic Rocks. *Bulletin of the Seismological Society of America*, 90(3):709–723, June 2000. doi: 10.1785/0119990142.

Cotton, F., Archuleta, R., and Causse, M. What is Sigma of the Stress Drop? *Seismological Research Letters*, 84(1):42–48, Jan. 2013. doi: 10.1785/0220120087.

Denolle, M. A. and Shearer, P. M. New perspectives on self-similarity for shallow thrust earthquakes. *Journal of Geophysical Research: Solid Earth*, 121(9):6533–6565, Sept. 2016. doi: 10.1002/2016jb013105.

Devin, E., Parker, G., Baltay, A., Nye, T., and Sahakian, V. Stress Drop in Ridgecrest Sequence Events from the Generalized Inversion Technique. In *AGU 2021 Fall Meeting*, page 13 – 17, December, New Orleans, 2021.

Dreger, D., Malagnini, L., Magana, J., and Taira, T. Comparing Finite-Source and Corner Frequency Based Stress Drop for the

Ridgecrest Sequence. In *AGU 2021 Fall Meeting*, page 13 – 17, December, New Orleans, 2021.

E.C.G.S. Workshop. Earthquake source physics on various scales, 2012. http://www.ecgs.lu/source2012.

Eshelby, J. The determination of the elastic field of an ellipsoidal inclusion, and related problems. In *Proceedings of the Royal Society of London Series A, 241(1226)*, 1957. doi: 10.1098/rspa.1957.0133.

Eulenfeld, T., Dahm, T., Heimann, S., and Wegler, U. Fast and Robust Earthquake Source Spectra and Moment Magnitudes from Envelope Inversion. *Bulletin of the Seismological Society of America*, 112(2):878–893, Nov. 2021. doi: 10.1785/0120210200.

Fan, W., Meng, H., Trugman, D., McGuire, J., and Cochran, E. Finite-source Attributes of M4 to 5.5 Ridgecrest, California Earthquakes. In *AGU 2022 Fall Meeting, 11-15*, December, Chicago, IL, abstract S15C-0209, 2022.

Gibowicz, S., Young, R., Talebi, S., and Rawlence, D. Source parameters of seismic events at the Underground Research Laboratory in Manitoba, Canada: Scaling relations for events with moment magnitude smaller than −2. *Bulletin of the Seismological Society of America*, 81(4):1157–1182, 1991. doi: 10.1785/BSSA0810041157.

Goertz-Allmann, B. P. and Edwards, B. Constraints on crustal attenuation and three-dimensional spatial distribution of stress drop in Switzerland. *Geophysical Journal International*, 196(1): 493–509, Oct. 2013. doi: 10.1093/gji/ggt384.

Goodfellow, S. D. and Young, R. P. A laboratory acoustic emission experiment under in situ conditions. *Geophysical Research Letters*, 41(10):3422–3430, May 2014. doi: 10.1002/2014gl059965.

Hanks, T. and Boore, D. Moment-magnitude relations in theory and practice. *Journal of Geophysical Research*, 89(B7), 1984. doi: 10.1029/JB089iB07p06229.

Hanks, T. C. Earthquake stress drops, ambient tectonic stresses and stresses that drive plate motions. *Pure and Applied Geophysics PAGEOPH*, 115(1–2):441–458, Jan. 1977. doi: 10.1007/bf01637120.

Hanks, T. C. and Kanamori, H. A moment magnitude scale. *Journal of Geophysical Research: Solid Earth*, 84(B5):2348–2350, May 1979. doi: 10.1029/jb084ib05p02348.

Hardebeck, J. L. and Aron, A. Earthquake Stress Drops and Inferred Fault Strength on the Hayward Fault, East San Francisco Bay, California. *Bulletin of the Seismological Society of America*, 99 (3):1801–1814, June 2009. doi: 10.1785/0120080242.

Heath, D. C., Wald, D. J., Worden, C. B., Thompson, E. M., and Smoczyk, G. M. A global hybrid VS30 map with a topographic slope–based default and regional map insets. *Earthquake Spectra*, 36(3):1570–1584, June 2020. doi: 10.1177/8755293020911137.

Huang, Y., Ellsworth, W. L., and Beroza, G. C. Stress drops of induced and tectonic earthquakes in the central United States are indistinguishable. *Science Advances*, 3(8), Aug. 2017. doi: 10.1126/sciadv.1700772.

Ide, S. and Beroza, G. C. Does apparent stress vary with earthquake size? *Geophysical Research Letters*, 28(17):3349–3352, Sept. 2001. doi: 10.1029/2001gl013106.

Ide, S., Beroza, G. C., Prejean, S. G., and Ellsworth, W. L. Apparent break in earthquake scaling due to path and site effects on deep borehole recordings. *Journal of Geophysical Research: Solid Earth*, 108(B5), May 2003. doi: 10.1029/2001jb001617.

Imanishi, K. and Ellsworth, W. L. *Source scaling relationships of microearthquakes at Parkfield, CA, determined using the SAFOD Pilot Hole Seismic Array*, page 81–90. American Geophysical Union, 2006. doi: 10.1029/170gm10.

Ji, C., Archuleta, R. J., and Wang, Y. Variability of Spectral Estimates of Stress Drop Reconciled by Radiated Energy. *Bulletin of the Seismological Society of America*, 112(4):1871–1885, June 2022. doi: 10.1785/0120210321.

Kaneko, Y. and Shearer, P. M. Variability of seismic source spectra, estimated stress drop, and radiated energy, derived from cohesive-zone models of symmetrical and asymmetrical circular and elliptical ruptures. *Journal of Geophysical Research: Solid Earth*, 120(2):1053–1079, Feb. 2015. doi: 10.1002/2014jb011642.

Kemna, K. B., Verdecchia, A., and Harrington, R. M. Spatio-Temporal Evolution of Earthquake Static Stress Drop Values in the 2016–2017 Central Italy Seismic Sequence. *Journal of Geophysical Research: Solid Earth*, 126(11), Nov. 2021. doi: 10.1029/2021jb022566.

Knudson, T., Shaw, B., and Beroza, G. Measuring Source Parameters With Filtered Peak S-Wave Amplitudes Using the Asymptotic Spectral Ratio Method. *Seismological Research Letters*, 94 (2B):1029–1314, Mar. 2023. doi: 10.1785/0220230054.

Ko, Y., Kuo, B., and Hung, S. Robust determination of earthquake source parameters and mantle attenuation. *Journal of Geophysical Research: Solid Earth*, 117(B4), Apr. 2012. doi: 10.1029/2011jb008759.

Kwiatek, G., Plenkers, K., and Dresen, G. Source Parameters of Picoseismicity Recorded at Mponeng Deep Gold Mine, South Africa: Implications for Scaling Relations. *Bulletin of the Seismological Society of America*, 101(6):2592–2608, Dec. 2011. doi: 10.1785/0120110094.

Lin, G., Shearer, P. M., Hauksson, E., and Thurber, C. H. A three-dimensional crustal seismic velocity model for southern California from a composite event method. *Journal of Geophysical Research: Solid Earth*, 112(B11), Nov. 2007. doi: 10.1029/2007jb004977.

Liu, M., Huang, Y., and Ritsema, J. Stress Drop Variation of Deep-Focus Earthquakes Based on Empirical Green's Functions. *Geophysical Research Letters*, 47(9), Apr. 2020. doi: 10.1029/2019gl086055.

Madariaga, R. Dynamics of an expanding circular fault. *Bulletin of the Seismological Society of America*, 66(3):639–666, June 1976. doi: 10.1785/bssa0660030639.

Malagnini, L., Mayeda, K., Nielsen, S., Yoo, S.-H., Munafo', I., Rawles, C., and Boschi, E. Scaling Transition in Earthquake Sources: A Possible Link Between Seismic and Laboratory Measurements. *Pure and Applied Geophysics*, 171(10):2685–2707, Dec. 2013. doi: 10.1007/s00024-013-0749-8.

Mayeda, K., Hofstetter, A., O'Boyle, J., and Walter, W. Stable and Transportable Regional Magnitudes Based on Coda-Derived Moment-Rate Spectra. *Bulletin of the Seismological Society of America*, 93(1):224–239, Feb. 2003. doi: 10.1785/0120020020.

Mayeda, K., Malagnini, L., and Walter, W. R. A new spectral ratio method using narrow band coda envelopes: Evidence for non-self-similarity in the Hector Mine sequence. *Geophysical Research Letters*, 34(11), June 2007. doi: 10.1029/2007gl030041.

McLaskey, G., Kilgore, B., Lockner, D., and Beeler, N. Laboratory generated M -6 earthquakes. *Pure and Applied Geophyics*, 171, 2014. doi: 10.1007/s00024-013-0772-9.

Morasca, P., Bindi, D., Mayeda, K., Roman-Nieves, J., Barno, J., Walter, W. R., and Spallarossa, D. Source scaling comparison and validation in Central Italy: data intensive direct Swaves versus the sparse data coda envelope methodology. *Geophysical Journal International*, 231(3):1573–1590, July 2022. doi: 10.1093/gji/ggac268.

Mori, J., Abercrombie, R. E., and Kanamori, H. Stress drops and radiated energies of aftershocks of the 1994 Northridge, California, earthquake. *Journal of Geophysical Research: Solid Earth*, 108(B11), Nov. 2003. doi: 10.1029/2001jb000474.

Nielsen, S., Spagnuolo, E., Smith, S. A. F., Violay, M., Di Toro, G., and Bistacchi, A. Scaling in natural and laboratory earthquakes. *Geophysical Research Letters*, 43(4):1504–1510, Feb. 2016. doi: 10.1002/2015gl067490.

Noda, H., Lapusta, N., and Kanamori, H. Comparison of average stress drop measures for ruptures with heterogeneous stress change and implications for earthquake physics. *Geophysical Journal International*, 193(3):1691–1712, Mar. 2013. doi: 10.1093/gji/ggt074.

Oye, V., Bungum, H., and Roth, M. Source Parameters and Scaling Relations for Mining-Related Seismicity within the Pyhäsalmi Ore Mine, Finland. *Bulletin of the Seismological Society of America*, 95(3):1011–1026, June 2005. doi: 10.1785/0120040170.

Pennington, C. N., Chen, X., Abercrombie, R. E., and Wu, Q. Cross Validation of Stress Drop Estimates and Interpretations for the 2011 Prague, OK, Earthquake Sequence Using Multiple Methods. *Journal of Geophysical Research: Solid Earth*, 126(3), Mar. 2021. doi: 10.1029/2020jb020888.

Rice, J. R. Heating and weakening of faults during earthquake slip. *Journal of Geophysical Research: Solid Earth*, 111(B5), May 2006. doi: 10.1029/2005jb004006.

Ruhl, C. J., Abercrombie, R. E., and Smith, K. D. Spatiotemporal Variation of Stress Drop During the 2008 Mogul, Nevada, Earthquake Swarm. *Journal of Geophysical Research: Solid Earth*, 122 (10):8163–8180, Oct. 2017. doi: 10.1002/2017jb014601.

Sato, T. and Hirasawa, T. Body wave spectra from propagating shear cracks. *Journal of Physics of the Earth*, 21(4):415–431, 1973. doi: 10.4294/jpe1952.21.415.

Satriano, C. SourceSpec – Earthquake source parameters from P- or S-wave displacement, 2022. spectra.DOI:10.5281/ZENODO.3688587.

SCEDC. Southern California Earthquake Data Center, 2013. doi: 10.7909/C3WD3XH1.

Sellers, E. J., Kataka, M. O., and Linzer, L. M. Source parameters of acoustic emission events and scaling with mining-induced seismicity. *Journal of Geophysical Research: Solid Earth*, 108(B9), Sept. 2003. doi: 10.1029/2001jb000670.

Shearer, P. M., Prieto, G. A., and Hauksson, E. Comprehensive analysis of earthquake source spectra in southern California. *Journal of Geophysical Research: Solid Earth*, 111(B6), June 2006. doi: 10.1029/2005jb003979.

Shearer, P. M., Abercrombie, R. E., Trugman, D. T., and Wang, W. Comparing EGF Methods for Estimating Corner Frequency and Stress Drop From P Wave Spectra. *Journal of Geophysical Research: Solid Earth*, 124(4):3966–3986, Apr. 2019. doi: 10.1029/2018jb016957.

Shearer, P. M., Abercrombie, R. E., and Trugman, D. T. Improved Stress Drop Estimates for M 1.5 to 4 Earthquakes in Southern California From 1996 to 2019. *Journal of Geophysical Research: Solid Earth*, 127(7), June 2022. doi: 10.1029/2022jb024243.

Shible, H., Hollender, F., Bindi, D., Traversa, P., Oth, A., Edwards, B., Klin, P., Kawase, H., Grendas, I., Castro, R. R., Theodoulidis, N., and Gueguen, P. GITEC: A Generalized Inversion Technique Benchmark. *Bulletin of the Seismological Society of America*, 112 (2):850–877, Feb. 2022. doi: 10.1785/0120210242.

Spottiswoode, S. and McGarr, A. Source parameters of tremors in a deep-level gold mine. *Bulletin of the Seismological Society of America*, 65(1):93–112, 1975. doi: 10.1785/BSSA0650010093.

Supino, M., Festa, G., and Zollo, A. A probabilistic method for the estimation of earthquake source parameters from spectral inversion: application to the 2016–2017 Central Italy seismic sequence. *Geophysical Journal International*, 218(2):988–1007, May 2019. doi: 10.1093/gji/ggz206.

Thatcher, W. and Hanks, T. C. Source parameters of southern California earthquakes. *Journal of Geophysical Research*, 78(35): 8547–8576, Dec. 1973. doi: 10.1029/jb078i035p08547.

Trugman, D. T. Stress-Drop and Source Scaling of the 2019 Ridgecrest, California, Earthquake Sequence. *Bulletin of the Seismological Society of America*, 110(4):1859–1871, May 2020. doi: 10.1785/0120200009.

Trugman, D. T. and Shearer, P. M. Application of an improved spectral decomposition method to examine earthquake source scaling in Southern California. *Journal of Geophysical Research: Solid Earth*, 122(4):2890–2910, Apr. 2017. doi: 10.1002/2017jb013971.

University of Nevada, Reno. NN Seismic Network [Data set], 1971. doi: 10.7914/SN/NN.

University of Nevada, Reno. SN Great Basin Network [Data set], 1992. doi: 10.7914/SN/SN.

Urbancic, T. and Young, R. Space-time variations in source parameters of mining-induced seismic events with M < 0. *Bulletin of the Seismological Society of America*, 83(2):378–397, 1993. doi: 10.1785/BSSA0830020378.

Urbancic, T. I., Trifu, C.-I., Mercer, R. A., Feustel, A. J., and Alexander, J. A. G. Automatic time-domain calculation of source parameters for the analysis of induced seismicity. *Bulletin of the Seismological Society of America*, 86(5):1627–1633, Oct. 1996. doi: 10.1785/bssa0860051627.

U.S. Geological Survey. United States National Strong-Motion Network, 1931. doi: 10.7914/SN/NP.

Vandevert, I., Shearer, P., and Fan, W. Earthquake Source Spectra Estimates Obtained from S-Wave Maximum Amplitudes: Application to the 2019 Ridgecrest Sequence. In *AGU 2022 Fall Meeting, 11-15*, December, Chicago, IL, abstract S25A-06, 2022.

Viesca, R. C. and Garagash, D. I. Ubiquitous weakening of faults due to thermal pressurization. *Nature Geoscience*, 8(11):875–879, Oct. 2015. doi: 10.1038/ngeo2554.

White, M. Ridgecrest 1D velocity model developed by Malcolm White, 2021. https://service.scedc.caltech.edu/ftp/stressdrop-ridgecrest/Ridgecrest_velocity_model.docx.

White, M. C. A., Fang, H., Catchings, R. D., Goldman, M. R., Steidl, J. H., and Ben-Zion, Y. Detailed traveltime tomography and seismic catalogue around the 2019 Mw7.1 Ridgecrest, California, earthquake using dense rapid-response seismic data. *Geophysical Journal International*, 227(1):204–227, June 2021. doi: 10.1093/gji/ggab224.

Yamada, T., Mori, J. J., Ide, S., Abercrombie, R. E., Kawakata, H., Nakatani, M., Iio, Y., and Ogasawara, H. Stress drops and radiated seismic energies of microearthquakes in a South African gold mine. *Journal of Geophysical Research: Solid Earth*, 112 (B3), Mar. 2007. doi: 10.1029/2006jb004553.

Yong, A., Martin, A., Stokoe, K., and Diehl, J. *ARRA-funded VS30 measurements using multi-technique approach at strong-motion stations in California and central-eastern United States*. 2013. doi: 10.3133/ofr20131102.

Yoshimitsu, N., Kawakata, H., and Takahashi, N. Magnitude −7 level earthquakes: A new lower limit of self-similarity in seismic scaling relationships. *Geophysical Research Letters*, 41(13): 4495–4502, July 2014. doi: 10.1002/2014gl060306.

Zhang, Q. and Lin, G. Three-dimensional Vp and Vp/Vs models in the Coso geothermal area, California: Seismic characterization of the magmatic system. *Journal of Geophysical Research: Solid Earth*, 119(6):4907–4922, June 2014. doi: 10.1002/2014jb010992.

REPORT
doi:10.26443/seismica.v3i1.981

History and activities of the European-Mediterranean Seismological Centre

Rémy Bossu *[1,2], **Florian Haslinger** [3], **Hélène Hébert** [2]

[1]EMSC, European–Mediterranean Seismological Centre, Arpajon, France, [2]CEA, DAM, DIF, Arpajon, France, [3]Swiss Seismological Service (SED) at ETH Zürich, Switzerland

Author contributions: *Conception*: RB. *Writing (original draft)*: RB. *Project administration*: RB, FH and HH. *Leadership*: RB, FH and HH. *Writing (review and editing)*: RB and FH.

Abstract The European-Mediterranean Seismological Centre (EMSC) provides rapid information on earthquakes and their effects, but does not operate seismic stations. It collects and merges parametric earthquake data from seismological agencies and networks around the world and collects earthquake observations from global earthquake eyewitnesses. Since its creation in 1975, it has developed strategies to complement earthquake monitoring activities of national agencies and coordinated its activities in Europe with its sister organisations ORFEUS and EFEHR as well as with global actors, while being part of the transformative EPOS initiative. The purpose of this article is to give a brief history of the EMSC and describe its activities, services and coordination mechanisms.

Production Editor:
Gareth Funning
Handling Editor:
Carmine Galasso
Copy & Layout Editor:
Théa Ragon

Received:
May 5, 2023
Accepted:
December 4, 2023
Published:
January 15, 2024

Introduction

The European-Mediterranean Seismological Centre (EMSC) has become one of the most important global earthquake information centres in the world over the last decades. While some of its activities are well known in the seismological community, the organisation itself, its history, structure and governance, its links with other European and global bodies, the way its services are organised and the basic principles that guide them have never been described in a single and open document and thus remain unclear to many actors in seismology and users of its services. The aim of this paper is to describe these different aspects of the EMSC and to illustrate how a regional non-profit non-governmental organisation can complement rapid public earthquake information in coordination with national actors thanks to a well-established and community-agreed policy. We also outline the current evolution of EMSC activities and the major overhaul of its processing system, and call for new networks to contribute data, as well as potential sponsors whose contributions are needed to maintain and further develop our activities and services.

EMSC brief history

In 1975, the European Seismological Commission (ESC), considering the level of seismic risk in the Euro-Mediterranean region, recommended the creation of the Centre Sismologique Euro-Mediterranéen (CSEM, or Euro-Mediterranean Seismological Centre, EMSC) to ”determine in near real time the epicenters of potentially damaging earthquakes” in this region, as well as the epicentral location of smaller magnitude earthquakes using data from existing monitoring networks (Mueller, 1980). This recommendation was supported by both IASPEI (International Association for Seismology and Physics of the Earth's Interior) and IUGG (International Union of Geodesy and Geophysics). In practice, it was a way of maintaining in Strasbourg the earthquake location activities of the Bureau Central International de Sismologie (BCIS, Rothé, 1981) which began publishing an instrumental catalogue in early 1900's and ceased to exist in 1975 in the Euro-Mediterranean region (Adams, 2002). EMSC practically started operating in 1976.

It may not be well known, but during the Cold War EMSC was instrumental in the global exchange of parametric data across the Iron Curtain. Direct telegraphic exchanges from some of the Warsaw Pact countries to the USA were restricted. The parametric data received at the EMSC by telex via the World Meteorological Organisation's Global Telecommunications System (WMO/GTS) were forwarded, still by telex, to the National Earthquake Information Center in Boulder, USA (NEIC was then part of the National Oceanic and Atmospheric Administration) and integrated into the Preliminary Determination of Epicenters (PDE) monthly bulletin, itself established in 1940. The Soviet bloc countries were aware of and pleased with this arrangement, which was seen as a way of solving a sensitive political problem (B. Presgrave, personal communication 2022).

*Corresponding author: bossu@emsc-csem.org

Data were usually sent to NEIC 2 to 3 times a week, thanks to Elie Peterschmitt and his staff at Louis Pasteur University in Strasbourg.

The EMSC, despite its lack of formal legal existence, continued to locate earthquakes in the Euro-Mediterranean region on an ad hoc basis until its founding meeting held in Strasbourg in December 1982 chaired by Jean Bonnin and attended by representatives from 8 countries (Belgium, Finland, France, Germany, Israel, Portugal, Switzerland, United-Kingdom) in addition to the ESC. Seismological institutes from 4 other countries (Albania, Italy, Spain, Yugoslavia) had expressed their support for this creation but did not attend the meeting. The statutes of the EMSC were presented in 1983 and officially registered in 1984 as a non-profit association under French law, a status that still exists today. The geographical area covered ranged from the Arctic in the north to the southern shores of the Mediterranean in the south, and from the Mid-Atlantic Ridge in the west to the Urals in the east. The aim was to rapidly locate earthquakes, improve data exchange, earthquake information and cooperation in the Euro-Mediterranean region.

In 1993, the agreement between the EMSC and its host Louis Pasteur University in Strasbourg was terminated and in 1994 the EMSC moved to the Laboratoire de Détection et de Géophysique (LDG) of the Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA) in Bruyères le Châtel, near Paris, its current location.

EMSC among the scientific bodies and actors

The EMSC operates under the auspices of the ESC (European Seismological Commission), the oldest regional commission of IASPEI (Adams, 2002). It coordinates its activities with its sister organisations in Europe, ORFEUS (Observatories & Research Facilities for European Seismology), a non-profit foundation for the coordination and promotion of digital broadband seismology in the Euro-Mediterranean area (Strollo, 2021), and more recently EFHER (European Facilities for Earthquake Hazard and Risk, Haslinger et al., 2022). Schematically, although operations, roles and responsibilities are different, in terms of services, EMSC is the European-Mediterranean version of NEIC (Hayes et al., 2011; Masse and Needham, 1989), while ORFEUS is that of IRIS-DMC (Incorporated Research Institutions for Seismology; Data Management Center. Smith, 1987; Hutko et al., 2017) now Earthscope.

The coordination between these three European organisations has been developed through a series of European funded projects for research infrastructures (e.g. Giardini et al., 2008), which in turn led to the establishment of EPOS (European Plate Observing System) as a European infrastructure for solid Earth sciences (Cocco et al., 2022). EMSC, ORFEUS and EFHER are jointly responsible for the seismology services within EPOS (Haslinger et al., 2022).

The EMSC was also involved in the now defunct UNESCO programme RELEMR (Reducing Earthquake Losses in Extended Mediterranean Region, https://en.unesco.org/disaster-risk-reduction/science-technology-resillience/REL) from the late 1990s to the mid-2010s to improve collaboration and data exchange with institutes around the Mediterranean. The bulletin exchanges established thanks to RELEMR significantly improved the availability of parametric data, adding readings from several hundred stations and in turn, the images of the seismicity in the region (Godey et al., 2006, 2013).

Membership, governance and funding

The EMSC has 3 types of membership, active members, key nodal members (a type of membership created in 1993 and introduced in 1994) and members by right. Active members are seismological institutes that participate in the activities of the EMSC and contribute to its objectives. Currently there are 66 of them from 54 countries (Table 1). Key Nodal Members are active members that provide specific support to the EMSC. Recognised Key Nodal Members are LDG (France) for hosting the EMSC, the GeoForschungsZentrum Potsdam (GFZ, Germany) for its key contribution to the EMSC services for global earthquake monitoring through its GEFON programme (Quinteros et al., 2021), the Istituto Nazionale di Geofisica e Vulcanologia (INGV, Italy) centres in Roma and Milan for thematic support on earthquake location methods and the AHEAD (Archive of Historical Earthquake Data) programme on European historical seismicity, respectively (Locati et al., 2014), and finally the Instituto Geografico Nacional (IGN, Spain) for maintaining a back-up website for EMSC members (www.ign.es/web/resources/sismologia/www/csem/fso.html). The ESC, the International Seismological Centre (ISC), NEIC/USGS and ORFEUS are members by right due to their international activities and cooperation with the EMSC.

The EMSC is governed by its annual General Assembly of members and advised by an Executive Council that consists of the President, three members elected by the General Assembly, representatives of the Key Nodal Members and the Secretary General. The Secretary General, who is responsible for day-to-day operations, administration, human resources and funding, is an employee of LDG, the host of the EMSC. The EMSC also benefits from the operational environment provided by LDG, which is responsible for informing the French authorities in case of earthquakes on the national mainland territory and operates the French Tsunami Warning Centre (Gailler et al., 2013; Roudil et al., 2013). LDG's support also includes the IT infrastructure of the EMSC and its maintenance.

Thanks to LDG hosting, the EMSC's expenses consist mainly of salaries and travel expenses of its staff, with minor allocations for other operational and administrative tasks. Funding comes from membership fees, participation in research projects (mainly European Union Framework Programmes), more recently EPOS, and sponsorship. A major challenge has been to maintain and improve services while being funded largely by soft money mainly dedicated to research. In 2020, the SCOR Foundation for Science offered a three-

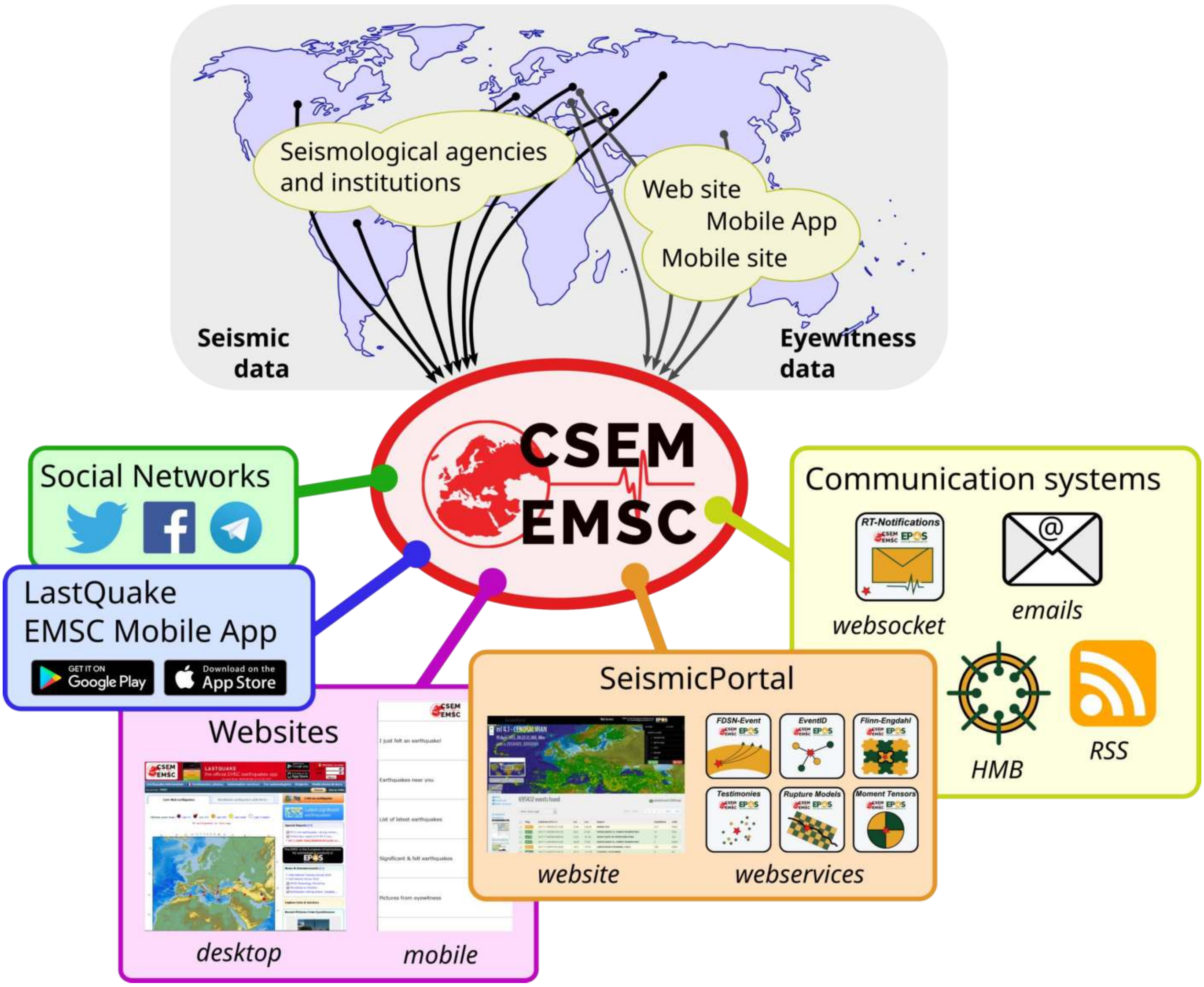


Figure 1 Schematic of EMSC services. Parametric data is collected from seismic networks to derive earthquake parameters, and eyewitness observations are collected through websites and the LastQuake smartphone app. Information is disseminated through various channels, including social networks and webservices.

year sponsorship to initiate a long overdue major upgrade of the service - the first in the last two decades - completed in June 2023. Sponsorship and financial donations remain an essential element of the EMSC's financial sustainability plan.

In addition to the Secretary General, there are currently 8 EMSC staff members comprising seismologists, IT experts, software developers and a sociologist. The size of the team has not changed recently and is unlikely to increase significantly due to the funding structure.

Roles and operation principles

The EMSC provides rapid information on earthquakes and their effects. It does not operate seismic stations. It merges seismic data, mainly parametric data (earthquake parameters, amplitudes, arrival times and CMTs) collected by network operators and crowdsourced ground truth data from eyewitnesses to provide services on a global scale with a focus on the Euro-Mediterranean region. (Figure 1; Table 2).

In contrast to many national seismological institutes, the EMSC has no legal mandate for earthquake information. Its scientific role is to provide redundancy and back-up to the authoritative national earthquake information services and to complement them, especially for earthquakes felt in several countries. Experience shows that redundancy and back-up may be needed after major earthquakes, as heavy traffic can bring down national earthquake information websites, hampering public communication and international data and information exchange. Merging seismic data can also improve earthquake information in border regions if bilateral exchange between neighboring countries is not optimal, or for offshore earthquakes. Complementarity of services is best illustrated by the online collection of macroseismic data, where collection at the national level optimizes the volume of data collected within national territories, but does not provide a complete picture when an earthquake is felt across borders. While methods exist to merge such geographically fragmented datasets (e.g., Van Noten et al., 2016), global-scale col-

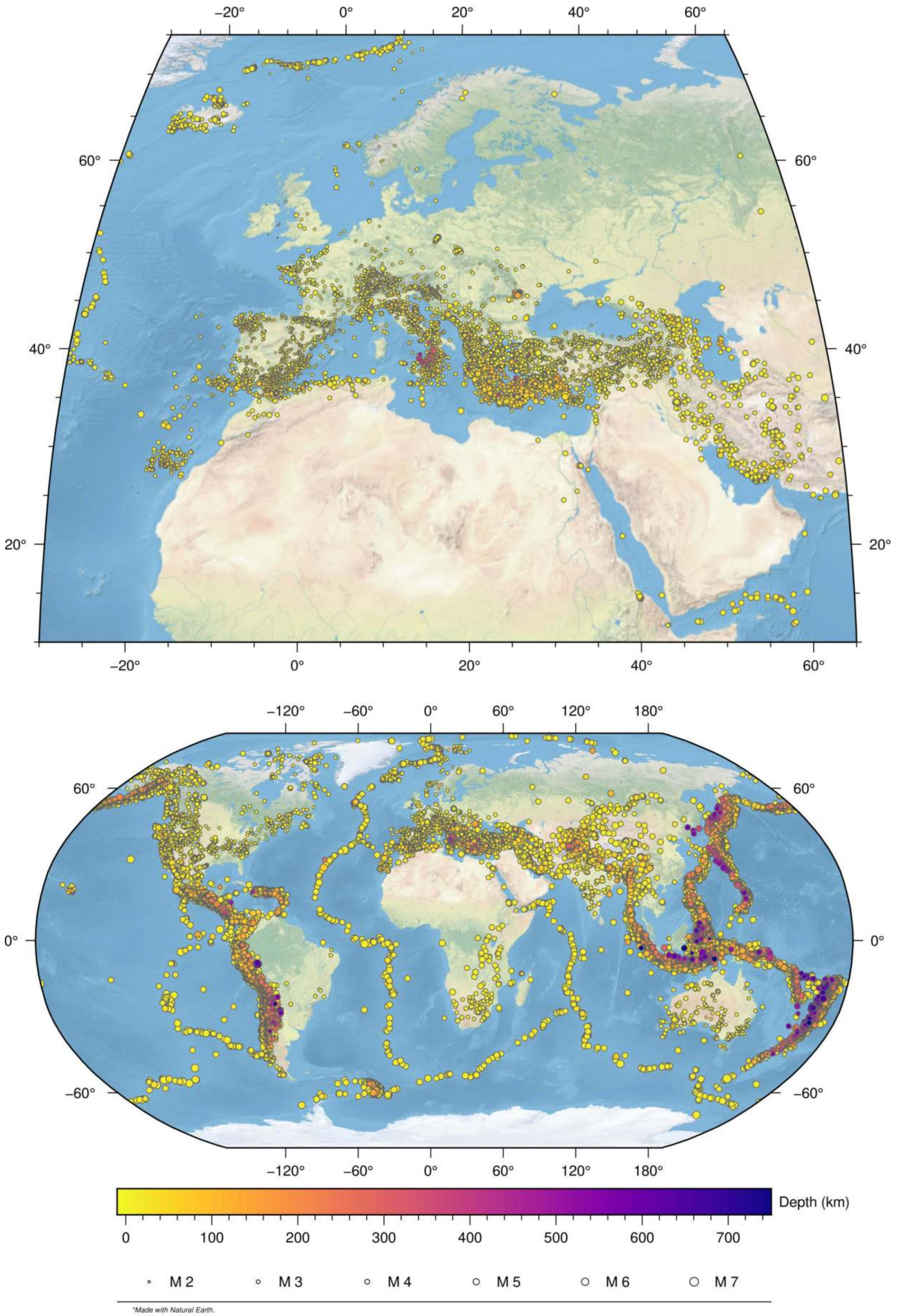


Figure 2 Geographical distribution of earthquakes reported in 2022: 22 148 earthquakes in the Euro-Mediterranean region (top) and 89 529 earthquakes on a global scale (bottom). Low-magnitude earthquakes are mainly reported in the Euro-Mediterranean region.

lection, like the one of the EMSC remains the fastest way to capture the full spatial distribution of impacts for such earthquakes.

One practical consequence of the lack of a legal mandate is that EMSC does not get involved in matters of national interest. In practice, it does not contact or develop projects with national civil protection services and media interviews on earthquake-related-issues are refused if they come from journalists in the affected countries. However, through its participation in the ARISTOTLE consortium, EMSC services contribute to the rapid earthquake impact assessment sent within 3 hours to the 24/7 Emergency Response Coordination Centre (ERCC), which is part of the EU Civil Protection Mechanism and coordinates the delivery of assistance to disaster-stricken countries (Michelini et al., 2020).

There are two basic principles for earthquake location at the EMSC, which were officially approved by the General Assembly in 2010 and described in Bossu and Mazet-Roux (2012). First, a provider can generally be trusted for earthquake information in the geographical area covered by its network, but its locations outside that area should not be reported unless they are consistently confirmed by another network. Application of this first principle implies that earthquake information can be maintained, at least for earthquakes large enough to be reported by several networks, even when information from the local network is not available. Second principle, relocations by the EMSC (obtained by merging the collected parametric data from the different contributors) should be limited to cases where a significant improvement in quality can be expected, or in other words, locations provided by data providers that are both reliable and accurate should be considered authoritative and published without change. In practice, a location is considered reliable if it can be reproduced with the associated data set of arrival times within its uncertainty range. It is considered accurate if it meets criteria related to the geometry and azimuthal distribution of reporting stations at short distances (up to 250 km, see details in Bossu and Mazet-Roux, 2012).

The implementation of these principles is more complex than described here, firstly because the system is fully dynamic, with new data constantly flowing in and manual observations replacing automatic ones. The implementation must also take into account the heterogeneity of network density and performance, and ensure the quality of information while avoiding missing significant earthquakes. For example, in a number of cases a moderate earthquake (M>4.5) was only reported by a local network within the boundaries of its network, while such an earthquake, given its magnitude, should have been reported by other networks, especially neighbouring ones. To cover such cases, a maximum magnitude is set for the network, above which an earthquake in its area of coverage will not be reported by the EMSC, unless confirmed by another network.

The presented approach of limiting the number of relocations performed by the EMSC is essential for public communication, where even slight discrepancies in earthquake locations between international organisations and national institutes can lead to misunderstandings and endanger public trust. It also implicitly recognises that the locations provided by national institutes are likely to be more accurate than those calculated by the EMSC using a similar dataset, thanks to their local knowledge and experience. By 2022, 85% of the 90,000 earthquake locations in the Euro-Mediterranean region publicly reported by the EMSC had been determined by data providers (Figure 2).

The situation for magnitude is more complex and magnitudes are not homogeneously determined. For small earthquakes, reported only by the local network, the magnitude is reported unchanged. For large earthquakes, the Mw provided notably by GFZ and NEIC is favoured. The main difficulty is for earthquakes 3<M<4.5, where available magnitude estimates are generally limited to ML (local magnitude) and often show large differences between different institutions. Whenever possible, the magnitude is recalculated using available amplitude measurements - if the definition and units are clearly defined - or using the EMSC instance of the SeisComp system (Weber et al., 2007). The final choice is then left to the seismologist performing the manual review.

Data contributors, data policy and data access

In 2022 there were 100 parametric data contributors, many of them EMSC members, representing a total of 8,130 seismic stations (Figure 3). The preferred data exchange tools are messaging systems, but despite our efforts to phase out email, it is still widely used because of its ease of setup. For 26 of these 100 contributors, earthquake location and magnitude are scanned from the institute's website when attempts to establish data exchange fail. In 2022, 4,871 focal mechanism and moment tensor solutions for 1,596 earthquakes were also collected from 12 different institutes. Finally, 249 000 felt reports representing the local level of shaking or damage were collected from earthquake eyewitnesses worldwide in 2022. The number varies as a function of seismicity and 250 000 have already been collected in the first 3 months of 2023 due to the earthquakes in Turkey.

All data collected is open, but no formal licensing of data and products has been finalised at this stage. This is a time-consuming process as it requires unanimity, certified by a signed document from each contributor. Thanks to the EPOS push, the aim is to apply the CC BY 4 licence (https://creativecommons.org) and eventually to meet the FAIR (Findability, Accessibility, Interoperability, and Reuse) principles.

Data can be accessed via the website (www.emsc-csem.org) or the earthquake portal. The website serves multiple audiences (public, scientists...) and provides fast information on earthquakes. It is more suitable for exploring individual events and recent activity, while the earthquake portal is aimed at researchers and provides access to larger datasets via web services (https://www.seismicportal.eu/webservices.html). Hosting the web services separately from the website limits the risk of slowdowns due to high traffic on the main site after

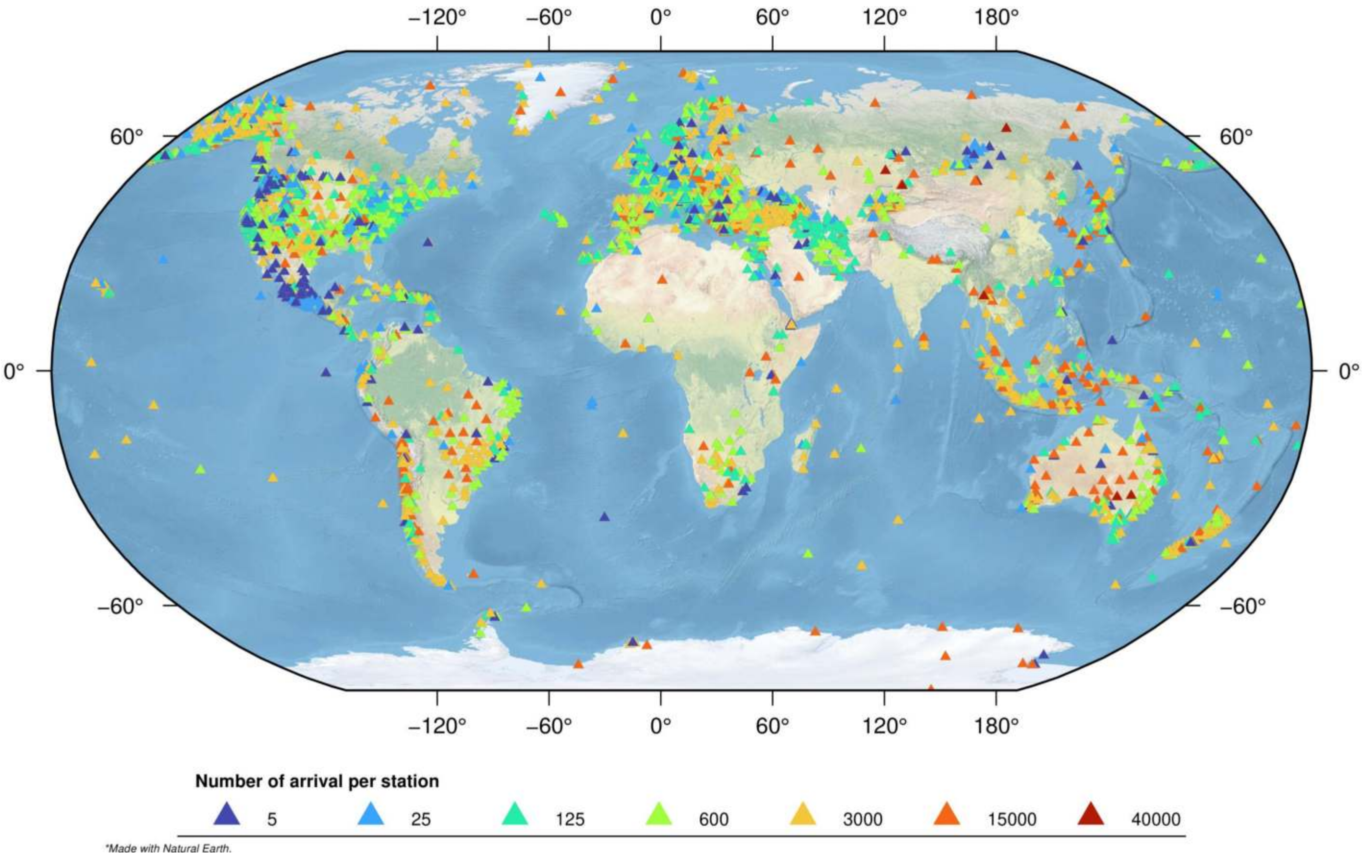


Figure 3 Locations of stations with reported arrival times in 2022, color depending on number of reported arrival times. Different organisations can pick phases from the same station due to open real-time waveform exchange. This means that EMSC can receive the same phase data for a station from multiple sources even if there is no parametric exchange between the station operator and the EMSC. However, parametric data exchange is essential for properly monitoring low magnitude local seismicity.

widely felt earthquakes. The FDSN event webservice (https://www.fdsn.org/webservices/fdsnws-event-1.2.pdf) is heavily used (average of 250 000 requests/day from 4 600 daily unique visitors). It was upgraded in 2023 and now has a limit of 20 000 events per request. The FDSN event service only publishes earthquake parameters once they have stabilised and so there is a typical delay before publication of a few to 20 minutes.

EMSC services

Although they are somewhat intertwined, the EMSC services can be schematically divided into 2 groups, one for the public and earthquake eyewitnesses, and one for the seismological community. The group of public activities, called LastQuake, aims to provide information about felt earthquakes and their effects. As it has been described in several publications, it is only outlined here.

LastQuake is a multi-component information and crowdsourcing system consisting of a smartphone application, a website for mobile devices and a Twitter bot (Bossu et al., 2018a, 2023). The eyewitness engagement strategy is based on crowdsourced detection, where felt earthquakes are detected not by seismic data, but by the online behavior of eyewitnesses immediately after they feel the shaking. Three types of crowdsourced detections are implemented at EMSC. Two of them reveal information-seeking behaviour, either by visiting our websites (Bossu et al., 2008, 2012, 2014) or by launching the LastQuake app (Bossu et al., 2018b), which generates a detectable and localizable change in the spatio-temporal characteristics of the traffic. The third, originally developed by Earle et al. (2012), monitors the rate of tweets (messages published on the microblogging site Twitter) containing the keyword "earthquake" in different languages, a rate that increases after a felt earthquake in a region where Twitter is popular, as eyewitnesses share their experiences.

Crowdsourced detections generally precede seismic locations and are typically available within 15 to 90 seconds of the earthquake. To be comprehensive, in 2022 these crowdsourced detections were supplemented by those independently performed by the Earthquake Network app, the first smartphone-based earthquake early warning (Finazzi, 2016). It detects felt earthquakes (Bossu et al., 2021) using the internal motion sensor of its users' smartphones. Crowdsourced detections are immediately published on the various components of the LastQuake system, and users are invited to confirm the existence of an earthquake by reporting their experience using a series of cartoons representing the 12 levels of the EMS 98 macroseismic scale (Grunthal, 1998). It initiates a rapid and massive collection of these re-

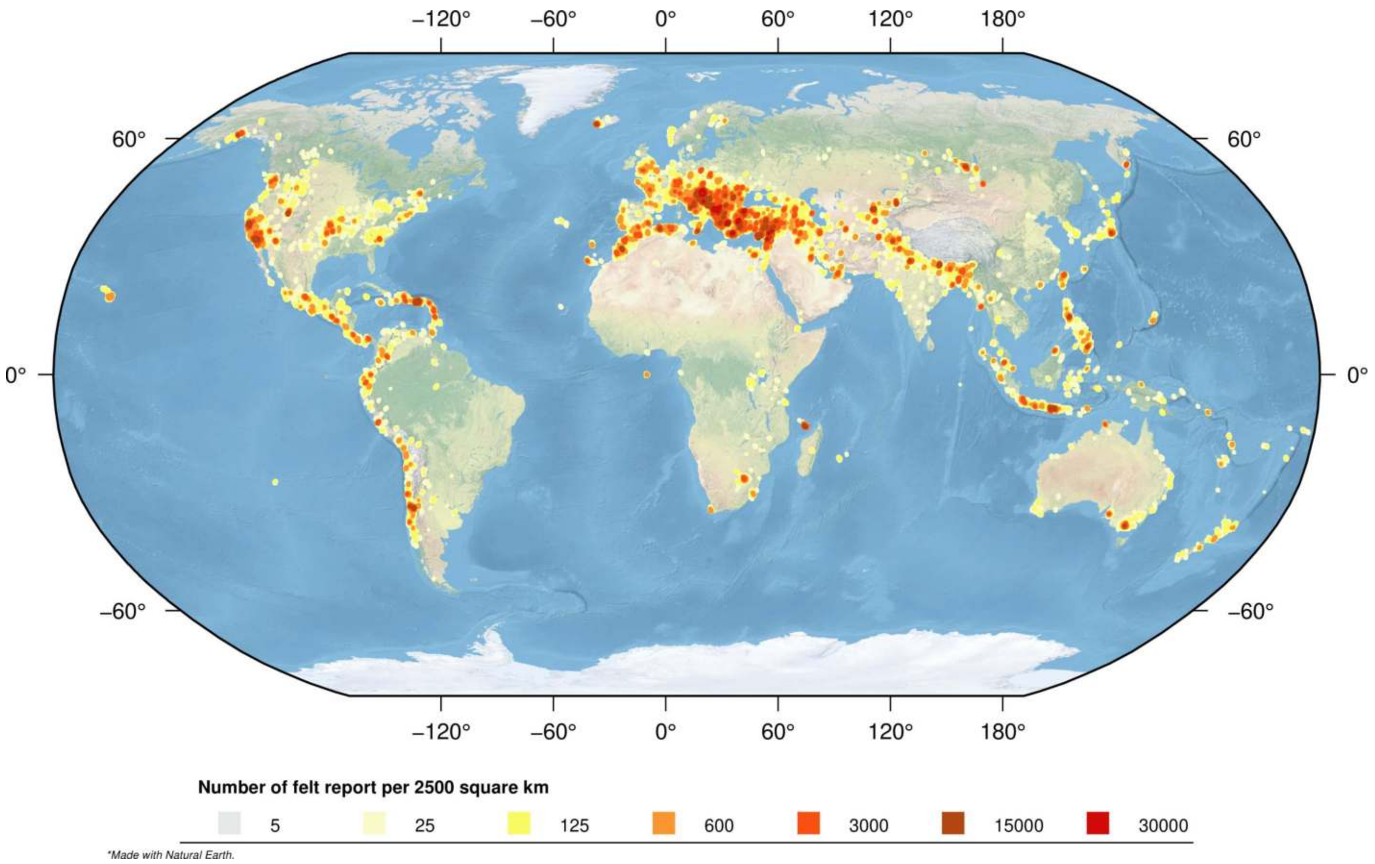


Figure 4 Geographical distribution of the density of the 2M felt reports crowdsourced up to April 18h 2023. The Europe-Mediterranean region is characterised by a high rate of crowdsourcing.

ports, called felt reports, with a median collection time of 10 minutes in 2022 (Figure 4). For example, more than 2,000 were collected within the first 15 minutes of the M7.8 2023, Kahramanmaras, Turkey earthquake. Felt reports are consistent with well calibrated "Did You Feel it?" (DYFI) responses (especially after a small correction of the bias for the high intensities, Wald et al., 1999; Quitoriano and Wald, 2020) as well as with independently and manually derived macroseismic datasets (Hough et al., 2016; Kouskouna et al., 2021; Bossu et al., 2015, 2017).

The determination and sharing of earthquake parameters have always been, and still is, the core service provided to the seismological community. Today, it deals exclusively with rapid determinations. However, a bulletin covering the European-Mediterranean region was produced for the period January 1998 to July 2012, which included data from 78 contributing networks from 53 countries and a total of 3,400 seismic stations. At the time, it significantly improved data availability in the region (Godey et al., 2006, 2013). However, due to funding difficulties and to avoid duplication with ISC activities, this activity has been discontinued. The bulletin is hosted at the ISC (https://doi.org/10.31905/EC1TT8WX) and data, metadata and local contacts have been transferred to the ISC to ensure long-term ingestion in its global bulletin. During this period, coordination with the ISC and NEIC was particularly close on issues such as the International Seismic Station Registry. The "For seismologists only" web page publishes the data sent by each contributor (earthquake parameters, moment tensors) as well as the parameters recalculated by the EMSC. It may contain several tens of locations for the same large earthquake, as determined by the different reporting networks. To limit misuse of the data, it contains a disclaimer pointing out the uncertain quality of the information, as many locations are fully automatic and outside the reporting networks. It is a popular webpage with network operators (20,000 accesses per day). The general public is invited to use the main page, which displays one set of parameters per earthquake. In the EMSC procedures, the data from the different networks are automatically merged. They are manually validated by EMSC staff during working hours, at least once a day during weekends and holidays, and for larger earthquakes in a more or less concentric scheme (M>5 in the Euro-Mediterranean region, M>6 in continental Asia and M>7 worldwide) by an on-call seismologist from our host institute in typically 20 minutes. Since July 2022, the Crowdseeded Seismic Location (CsLoc) method, which combines crowdsourced detections and seismic data analysis for fast (60-90s) and reliable locations of felt earthquakes (Steed et al., 2019; Bondár et al., 2020), has been fully implemented.

The only service restricted to members and accessible by login is the results of rapid impact assessment by the tool named EQIA (Earthquake Qualitative Impact Assessment, Julien-Laferrière, 2019; Guérin-Marthe et al., 2021). It offers heads-up on the scale of damage. However, such result is considered vulnerable to misinterpretation and over-interpretation by laypersons and journalists due to the inherent uncertainties of

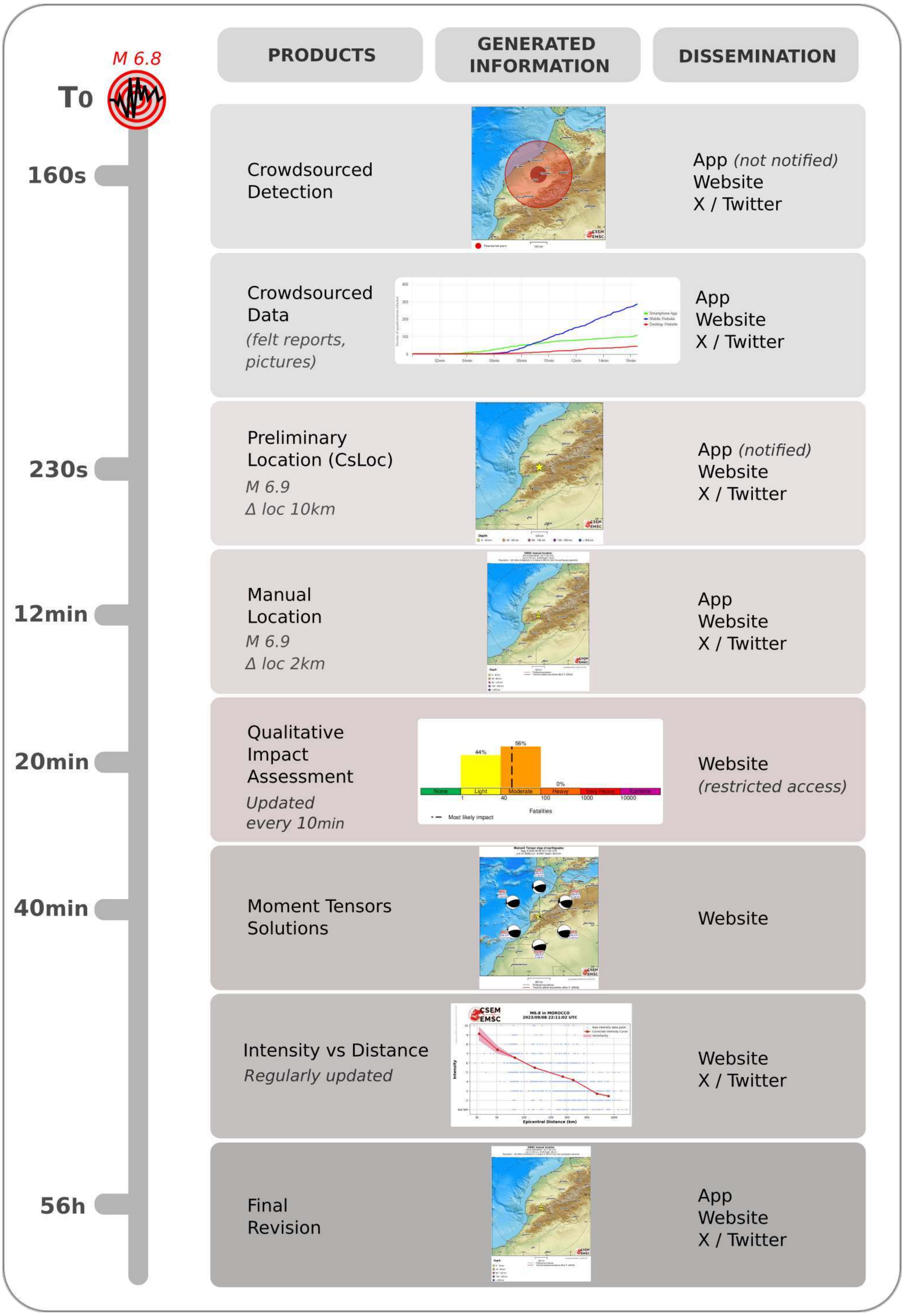


Figure 5 Abbreviated timeline -relative to origin time- of the main EMSC product releases and updates as well as their distribution channels for the 8 September 2023 M 6.8 Morocco earthquake.

such estimates and access is restricted to identified end-users (Bossu et al., 2015). Figure 5 shows an abbreviated timeline -relative to origin time- of the main EMSC product releases and updates as well as their distribution channels for the 8 September 2023 M 6.8 Morocco earthquake.

Current evolutions

There have been 2 major developments over the last few years. The first is the complete refactoring of both the back-end (processing part) and front-end (websites, smartphone app, etc.) systems, made possible thanks to the support of the SCOR for Science Foundation. The second concerns new methods related to rapid impact

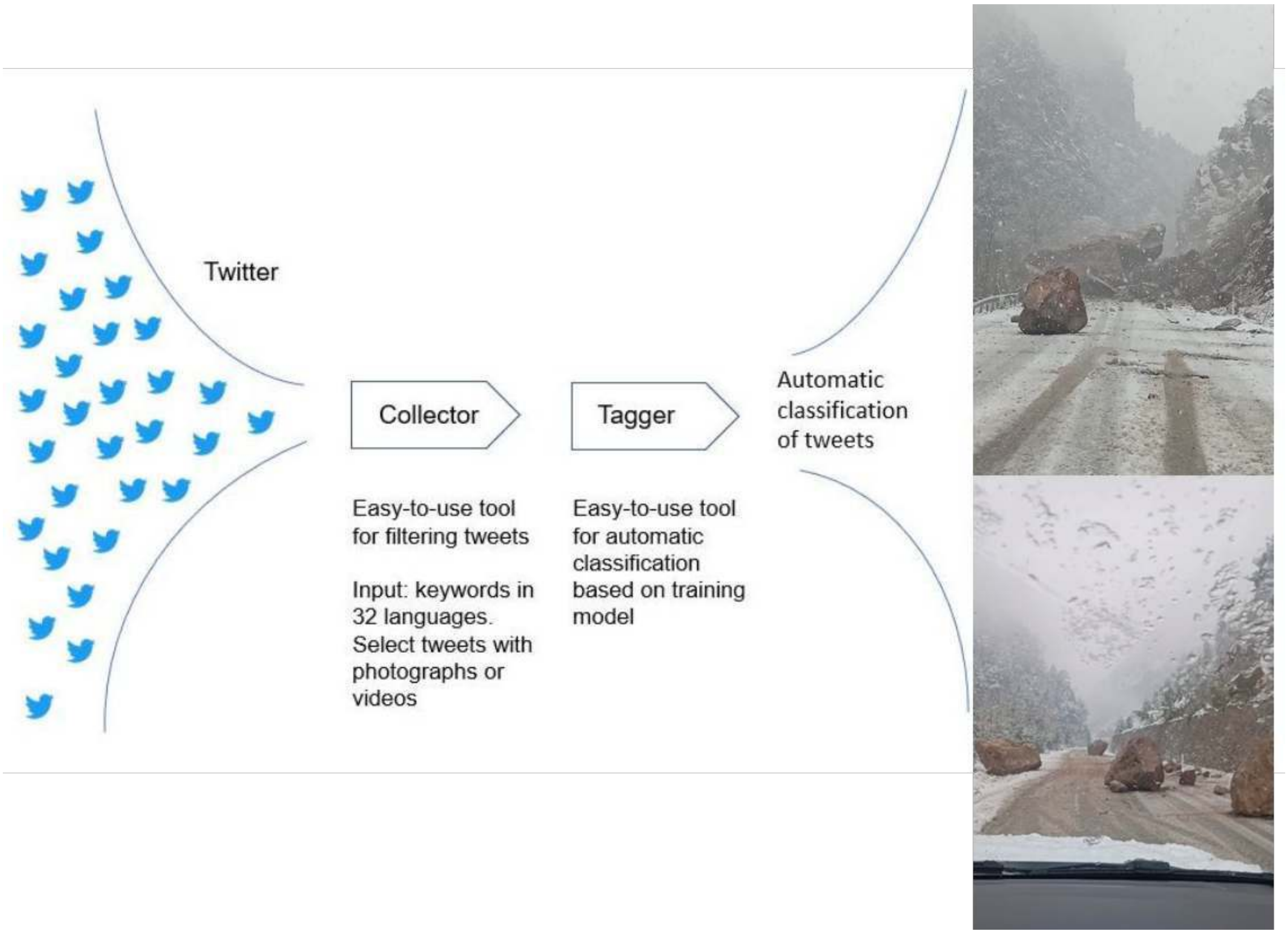


Figure 6 Schematic of the Global Landslide Detector (GLD), developed in collaboration with the Qatar Computing Research Institute and the British Geological Survey Landslide Team. Tweets containing both an image and a landslide-related word in different languages are collected. Images not related to landslides are automatically rejected. The 2 images on the right were collected 12 hours after the M7.8 Kahramanmaras earthquake in Turkey.

assessment, some of which optimise the use of felt reports for the calculation of shaking maps and damage assessment, and other harvesting of information from social media for the detection of landslides.

The refactoring of the systems is the first of its kind and was long overdue. It started with a new mobile website in 2020, followed by a new version of the Twitter bot (Twitter is now called X) in February 2022 (Bossu et al., 2023). The new version of the smartphone app is currently being tested and a new desktop website was launched at the end of June 2023. The main change concerns the backend and the processing of the seismic data, including in particular a new data model, a modular structure and the implementation of the iLoc location algorithm, particularly suitable for unbalanced networks and with more accurate formal uncertainty estimates (Bondár et al., 2018). It was originally developed at the ISC, where its implementation has resulted in consistent locations improvements (Bondár and Storchak, 2011). The new system and associated website for desktops allow better crediting of data contributors and different types of contributions (e.g. phase picking, station operators...). By adding a third digit to the earthquake locations, the grid patterns visible on the highly zoomed map of the earthquake sequence accompanying the Cumbia Vieja volcano on the island of La Palma in 2021, which led to rumours and conspiracy theories, will not be repeated (Fallou et al., 2022). Only minor adjustments have been made to the new seismic data processing system since its release in June 2023. So far, the main use of the felt reports has been limited to purely data-driven products such as earthquake impact maps and intensity vs distance curves. This is now evolving rapidly. Quitoriano and Wald (2022) developed a methodology to incorporate them into ShakeMap products, resulting in a lower level of uncertainty. Böse et al. (2021) apply the Finite-Fault Rupture Detector (FinDer) algorithm, which typically requires real-time ground motion observations from a dense seismic network operated in the vicinity of the earthquake (Böse et al., 2012), to felt reports to compute line-source models. The system has been in operation for the last 18 months and the results already appear promising when using the first 10 or 20 min of felt reports, but they still need to be fully evaluated. The inclusion of both felt reports and, for larger earthquakes, an early finite rupture model, could significantly speed up the production of reliable Shakemaps for global earthquakes, and in turn reduce the uncertainties of the impact models derived from ShakeMap. Recently, Lilienkamp et al. (2023) developed a data-driven approach that bypasses the computation of ShakeMap and is completely independent of seismic data to discriminate high-impact from low-impact earthquakes globally based only on felt reports available within the first 10 minutes. It is a first step and could evolve into a traffic light system by adding additional crowdsourced data. However, it can already correctly classify a significant proportion (39%) of low-impact events with high confidence and then quickly and reliably rule out the need for emergency response. Still related to the use of crowdsourced data, Contreras et al. (2022) performed a sentiment and topic analysis on the comments of users providing felt reports. As hypothesised, negative polarity in the comments is associated with higher intensities, while positive polarity prevails in those associated with the lowest intensities.

Finally, a prototype called the "Global Landslide Detector" is available online (https://landslide-aidr.qcri.org/service.php#home). It collects tweets (messages published on Twitter) containing both the keyword "landslide" and related words in different languages, as well as an image. A trained AI engine rejects the images not related to landslides (more than 99% of the collected tweets, Pennington et al., 2022; Ofli et al., 2022, Figure 6). Initiated by the EMSC to detect triggered landslides, which can significantly hamper rescue operations by blocking roads, the project was expanded to detect and document all types of landslides. A landslide was detected 12 hours after the M7.8 earthquake in Kahramanmaraş, Turkey (Figure 6). GLD's operations are currently affected by Twitter's data access restrictions. These developments aim to improve the ability to quickly and reliably assess the impact of global earthquakes.

Concluding remarks

The EMSC is a non-profit organisation created by the seismological community to provide it with rapid information on earthquakes and their effects, with a portfolio of services complementary to those of the national institutes. It has benefited from numerous European research projects to fund the development of its services and to pioneer citizen seismology, and is now implementing a sustainability plan thanks to its participation in long-term initiatives such as EPOS or ARISTOTLE, but also thanks to private donations and sponsorships. It has an open data policy and aims to improve its dissemination services in the coming years. Finally, this paper is also an opportunity to call on network operators to consider sharing their parametric data.

Acknowledgments

The EMSC as an organisation and its services are the result of the long-term support and cooperation of many people and institutions over the last 20 years. Firstly, the members for their support and active participation in setting policy, past and present members of the Executive Council for their guidance, with a special mention to past EMSC President Chris Browitt, and without forgetting the data contributors. They are all warmly thanked and should be given credit for the achievements of the EMSC. The services have been developed thanks to the dedicated work of past and present staff. The authors would like to thank Michel Cara and Bruce Presgrave for their insightful discussions on the history of the EMSC as well as the SCOR Foundation for Science for enabling the technical upgrading of the services. Reviews by A. Michelini, J. Schweitzer and 2 anonymous reviewers have helped to improve the quality of the manuscript.

Competing interests

The authors declare no competing interests

Table 1 (continued)

Key Nodal Members	
Laboratoire de Détection et de Géophysique (LDG)	France
GeoForschungsZentrum (GFZ)	Germany
Istituto Nazionale di Geofisica e Vulcanologia (INGV)	Italy, Roma
Istituto Nazionale di Geofisica e Vulcanologia (INGV)	Italy, Milano
Instituto Geografico Nacional (IGN)	Spain
Active Members	
Institute of Geosciences, Polytechnic University of Tirana (IGEO)	Albania
Centre de Recherche en Astronomie, Astrophysique et Geophysique (CRAAG)	Algeria
National Survey for Seismic Protection (NSSP)	Armenia
GeoSphere Austria	Austria
Republican Seismic Survey Center of Azerbaijan National Academy of Sciences (RSSC)	Azerbaijan
Center of Geophysical Monitoring (CGM)	Belarus
Royal Observatory of Belgium (ORB/ROB)	Belgium
Republic Hydrometeorological Institute (RHI)	Bosnia-Herzegovina
Federal Meteorological Institute (FMI)	Bosnia-Herzegovina
National Institute in Geophysics, Geodesy and Geography - BAS	Bulgaria
Croatian Seismological Survey (CSS)	Croatia
Geological Survey Department (GSD)	Cyprus
Institute of Physics of the Earth, Brno (IPE)	Czech Republic
Geophysical Institute of the Academy of Sciences (GFU)	Czech Republic
Geological Survey of Denmark and Greenland (GEUS)	Denmark
Observatoire Geophysique d'Arta (CERD)	Djibouti
National Research Institute of Astronomy and Geophysics (NRIAG)	Egypt
Institute of Seismology, University of Helsinki (ISUH)	Finland
Bureau Central Sismologique Francais (BCSF)	France
ISTerre, Institut des Sciences de la Terre	France
Seismic Monitoring Centre of Georgia (SMC)	Georgia
Federal Institute for Geosciences and Natural Resources (BGR)	Germany
National Observatory of Athens (NOA)	Greece
University of Thessaloniki (AUTH)	Greece
Institute of Engineering Seismology and Earthquake Engineering (ITSAK)	Greece
Laboratory of Seismology, University of Patras	Greece
Kövesligethy Radó Seismological Observatory	Hungary
Icelandic Meteorological Office (IMO)	Iceland
Dublin Institute for Advanced Studies (DIAS)	Ireland
Geological Survey of Israel (GSI)	Israel
National Data Center (NDC) of Israel, Soreq Nuclear Research Center	Israel
Istituto Nazionale di Oceanografia e Geofisica Sperimentale (OGS)	Italy
Jordan Seismological Observatory	Jordan
Seismological Institute of Kosovo	Kosovo
Geophysics Centre at Bhannes (SGB)	Lebanon
Libyan Center for Remote Sensing and Space Science (LCRSSS)	Libya
European Center for Geodynamics and Seismology (ECGS)	Luxembourg
Seismological Observatory	North Macedonia
Department, University of Malta (UM)	Malta
Institute of Geology and Seismology	Moldova
Direction de l'Environnement	Monaco
Institute of Hydrometeorology and Seismology (MSO)	Montenegro
Centre National pour la Recherche Scientifique et Technique (CNRST)	Morocco
Département des Sciences de la Terre	Morocco
University of Bergen (BER)	Norway

NORSAR	Norway
Institute of Geophysics, Polish Academy of Sciences (IGPAS)	Poland
Instituto de Meteorologia (IMP)	Portugal
Universidade de Evora	Portugal
Faculdade de Ciências da Universidade de Lisboa	Portugal
National Institute for Earth Physics (NIEP)	Romania
Geophysical Survey of the Russian Academy of Sciences (GSRAS)	Russia
Seismological Survey of Serbia (SSS)	Serbia
Earth Science Institute, SAS, Department of Seismology	Slovakia
Agencija Republike Slovenije za okolje (ARSO)	Slovenia
Institut Cartografic i Geologic de Catalunya (ICGC)	Spain
Swedish National Seismic Network (SNSN)	Sweden
Schweizerischer Erdbebendienst (ETH)	Switzerland
Royal Netherlands Meteorological Institute (KNMI)	The Netherlands
Institut National de la Météorologie (INMT)	Tunisia
Disaster and Emergency Management Presidency, Earthquake Department (ERD)	Turkey
Kandilli Observatory and Earthquake Research Institute (KOERI)	Turkey
Main Centre for Special Monitoring (MCSM)	Ukraine
Dubai Municipality Seismic Network	United Arab Emirates
British Geological Survey (BGS)	United Kingdom
National Seismological Observatory Centre (NSOC)	Yemen
Members by right	
European Seismological Commission (ESC)	
Observatories and Research Facilities for EUropean Seismology (ORFEUS)	
International Seismological centre (ISC)	
U.S. Geological Survey (USGS)	United States

Table 1 List of member institutions in January 2023.

Table 2 (continued)

Institute	Country/Region	Exchange tool	Parametric data	MT
Institute of Geosciences, Polytechnic University of Tirana (IGEO)	Albania	Email	L P A	MT
Centre de Recherche en Astronomie, Astrophysique et Géophysique (CRAAG)	Algeria	Web	L	
Instituto Nacional de Prevencion Sismica (INPRES) (NSNA)	Argentina	Web	L	
National Survey of Seismic Protection (NSSP)	Armenia	Email	L P A	
Geoscience Australia, Canberra, ACT, Australia (AUST)	Australia	Mail	L P A	
Geosphere Austria (GBA)	Austria	Email	L P	
Republican Seismic Survey Center or Azerbaijan National Academy of Sciences (RSSC)	Azerbaijan	Email	L P A	
Royal Observatory of Belgium (UCC)	Belgium	Email	L P A	
Rede Sismografica Brasileira (RSBR)	Brazil	Web	L P A	
National Institute in Geophysics, Geodesy and Geography - BAS (SOF)	Bulgaria	Email	L P A	
Canadian National Seismic Network (CNSN) BB stations (CN)	Canada	Web	L	
Departamento de Geofisica, Universidad de Chile (CSN)	Chile	Email	L	
Seccion de Sismologia, Univ. de Costa Rica, San Jose, Costa Rica (UCR)	Costa Rica	Web	L	
Seismological Survey,University of Zagreb (ZAG)	Croatia	Email	L P	
Servicio Sismologico Nacional de Cuba (CENAIS) (SSNC)	Cuba	Web	L	
Geological Survey Department (GSD)	Cyprus	Email	L P A	
Geophysical Institute of the Academy of Sciences (GFU)	Czech Rep.	Email	L P	
Institute of Physics of the Earth (IPEC)	Czech Rep.	Email	L P A	

Table 2 (continued)

Universidad Autonoma de Santo Domingo (UASD)	Dominican Rep.	Web	L	
Escuela Politecnica Nacional, Quito, Ecuador (QUI)	Ecuador	Web	L	
National Research Institute of Astronomy and Geophysics (NRIAG)	Egypt	Email	L P A	
Servicio Nacional de Estudios Territoriales (SNET)	El Salvador	Web	L	
Laboratoire de Detection et de Geophysique (LDG)	France	Email	L P A	
Institut de Physique du Globe de Paris (IPGP)	France	Email		DC
Géoazur (Université Cote d'Azur, IRD, CNRS, Observatoire de la Cote d'Azur) (OCA)	France	Email	L P A	DC
Réseau National de Surveillance Sismique (ReNaSS)	France	Web	L P A	
Seismic Monitoring Centre of Georgia (TIF)	Georgia	Email	L P	
Bundesanstalt fur Geowissenschaften und Rohstoffe, German Regional Seismograph Network (BGR)	Germany	Email	L P A	
GeoForschungsZentrum (GFZ)	Germany	HMB	L P A	MT
Landsamt fur Geologie, Rohstoffe und Bergbau (LED)	Germany	Email	L P	
National Observatory of Athens, Geodynamic Institute (NOA)	Greece	Email	L P A	MT
Aristotle University of Thessaloniki, Department of Geophysics (THE)	Greece	Email	L P A	
University Of Athens (UOA)	Greece	Email		MT
University of Patras Seismological Laboratory (UPSL)	Greece	Email		MT
Observatoire Volcanologique et Sismologique de Guadeloupe (OVSG - IPGP) (OVSG)	Guadeloupe	Web	L P A	
URGeo, Geoazur (Universite Cote d'Azur, IRD, CNRS, Observatoire de la Cote d'Azur) (AYIT)	Haiti	Email	L P A	
MTA CSFK GGI Kovesligethy Rado Seismological Observatory (BUD)	Hungary	Email	L P A	
Department of Geophysics, Icelandic Meteorological Office (IMO)	Iceland	Web	L	
India Meteorological Department, New Delhi, India (NDI)	India	Web	L	
Badan Meteorologi, Klimatologi dan Geofisika (BMKG)	Indonesia	Web	L	
Institute of Geophysics, University of Tehran (IGUT)	Iran	Email	L P A	
International Institute for Earthquake Engineering and Seismology (IIEES)	Iran	Email	L	
Irish National Seismic Network (INSN)	Ireland	Email	L P A	
Geological Survey of Israel, Seismology Division (GSI)	Israel	Email	L P	
Instituto Nazionale di Geofisica e Vulcanologia (INGV)	Italy	Email	L P A	MT
Instituto Nazionale di Oceanografia e Geofisica Sperimentale - OGS (OGS)	Italy	Email	L P A	
Kazakhstan National Data Center (KNDC)	Kazakhstan	Email	L P A	
Korean Meteorological Administration (SEO)	S. Korea	Web	L	
Kyrgyz Institute of Seismology (KIS)	Kyrgyzstan	Email	L P A	
National Center for Geophysical Research (GRAL)	Lebanon	Email	L P A	
Malaysian Meteorological Department (MMD)	Malaysia	HMB	L P A	
Malta Seismic Network, Seismic Monitoring and Research Unit (SMRU), University of Malta (MLT)	Malta	Email	L P A	
Observatoire Volcanologique et Sismologique de Martinique (OVSM - IPGP) (OVSM)	Martinique	Web	L P	
Servicio Sismologico Nacional, Instituto de Geofisica, UNAM (UNM)	Mexico	Web	L	
Institute of Geophysics and Geology (MOLD)	Moldova	Email	L P A	
Montenegro Seismological Observatory (MSO)	Montenegro	Email	L P A	
Centre National de la Recherche Scientifique et Technique (CNRST)	Morocco	Email	L P	
National Seismological Centre, Department of Mines and Geology (NSC)	Nepal	Web	L	
Koninklijk Nederlands Meteorologish Instituut (KNMI)	Netherlands	Web	L	
Geonet, GNS science (GNS)	New Zealand	Web	L P	

Table 2 (continued)

Instituto Nicaraguense de Estudios Territoriales (INET)	Nicaragua	Web	L	
Seismological Observatory (SKO)	N. Macedonia	Email	L P A	
University of Bergen (BER)	Norway	Email	L P A	
NORSAR	Norway	Email	L P A	
Centre Polynésien de Préventions des Tsunamis (CPPT)	Pamatai	Email		MT
Universidad de Panama (IGC)	Panama	Web	L	
Instituto Geofisico del Peru (LIM)	Peru	Web	L	
Philippine Inst. of Volcanology and Seismology, Quezon City, Philippines (PIVS)	Philippines	Web	L	
Instituto Portugues do Mar e da Atmosfera (IPMA)	Portugal	Email	L P A	
Instituto Portugues do Mar e Atmosfera (PDA)	Portugal	Email	L P A	
Puerto Rico Seismic Network (PRSN) and Puerto Rico Strong Motion Program (PRSMP), University of Puerto Rico at Mayaguez (PR)	Puerto Rico	PDL	L P A	
National Institute for Earth Physics (NIEP)	Romania	Email	L P	
Geophysical Survey of the Russian Academy of Sciences (GSRAS)	Russia	Email	L P	
Seismological Survey of Serbia (SSS)	Serbia	Email	L P A	
Agencija Republike Slovenije za okolje, Seismological Office (LJU)	Slovenia	Email	L P	
South African Seismological Network (SASN)	South Africa	Web	L	
Instituto Cartografic i Geologic de Catalunya (ICGC)	Spain	Email	L P A	
Instituto Geografico Nacional (IGN)	Spain	Email	L P A	
Swiss Seismological Service (ETHZ)	Switzerland	Email	L P A	
Central Weather Bureau (CWB)	Taiwan	Email	L P	
Thailand Seismological Bureau (TSB)	Thailand	Web	L	
University of the West Indies, St. Augustine, Trinidad (TRN)	Trinidad and Tobago	Email	L	
Institut National de Meteorologie (INMT)	Tunisia	Email	L P A	
Disaster and Emergency Management Presidency, Earthquake Department (AFAD)	Turkey	Email	L P A	MT
Kandilli Observatory and Earthquake Research Institute (KOERI)	Turkey	Email	L P	MT
Carpathian Seismological Department, Ukraine Academy of Science (LVV)	Ukraine	Email	L P	
Ukrainian NDC, Main Center of Special Monitoring (MCSM)	Ukraine	Email	L P A	
British Geological Survey (BGS)	United Kingdom	Email	L P A	
Alaska Regional Network, University of Alaska-Fairbanks (AK)	US	PDL	L P A	
Alaska Tsunami Warning Seismic System, West Coast and Alaska Tsunami Warning Center (AT)	US	PDL	L P A	
Alaska Volcano Observatory, USGS - Anchorage, University of Alaska, Geophysical Institute (AV)	US	PDL	L P A	
Southern California Seismic Network, California Institute of Technology / USGS - Pasadena (SCSN)	US	PDL	L P A	
Hawaiian Volcano Observatory Network, Hawaiian Volcano Observatory (HV)	US	PDL	L P A	
Montana Regional Seismic Network, Montana Bureau of Mines and Geology (MB)	US	PDL	L P A	
USGS Northern California Regional Network, USGS-Menlo Park, California (NC)	US	PDL	L P A	
National Earthquake Information Center, U.S. Geological Survey (NEIC)	US	PDL	L P A	MT
Cooperative New Madrid Seismic Network, St. Louis University and University of Memphis (NM)	US	PDL	L P A	
Western Great Basin/Eastern Sierra Nevada, University of Nevada, Reno (NN)	US	PDL	L P A	
Oklahoma Seismic Network, University of Oklahoma (OK)	US	PDL	L P A	
Pacific Tsunami Warning Seismic System, Pacific Tsunami Warning Center, Ewa Beach, Hawaii (PT)	US	PDL	L P A	

Table 2 (continued)

Southeastern Appalachian Cooperative Seismic Network, Virginia Tech, University of Memphis, Tennessee Valley Authority, and University of North Carolina (SE)	US	PDL	L P A	
Bureau of Economic Geology, The University of Texas at Austin (BEG UTEXAS) (TX)	US	PDL	L P A	
University of Utah Regional Network, University of Utah Seismograph Stations (UU)	US	PDL	L P A	
Pacific Northwest Regional Seismic Network, University of Washington, Seattle (UW)	US	PDL	L P A	
Global Centroid-Moment-Tensor (GCMT)	US	Email		MT

Table 2 List of data contributors in 2022 for both earthquake parametric data and moment tensors. Contributions are sent via email or messaging systems (PDL or HMB). In some cases, they come from scrapping institutions' websites (Web). Parametric data contains at least locations and magnitude (L). They generally contains picks (P) and amplitudes (A).

References

Adams, R. International seismology. *International Handbook of Earthquake and Engineering Seismology*, page 29–37, 2002. doi: 10.1016/s0074-6142(02)80207-8.

Bondár, I. and Storchak, D. Improved location procedures at the International Seismological Centre: Improved location procedures at the ISC. *Geophysical Journal International*, 186(3): 1220–1244, July 2011. doi: 10.1111/j.1365-246x.2011.05107.x.

Bondár, I., Cs. Czanik, M., Gráczer, Z., Wéber, Z., and Group, A. Relocation of Seismicity in the Pannonian Basin Using a Global 3D Velocity Model. *Seismological Research Letters*, Oct. 2018. doi: 10.1785/0220180143.

Bondár, I., Steed, R., Roch, J., Bossu, R., Heinloo, A., Saul, J., and Strollo, A. Accurate Locations of Felt Earthquakes Using Crowdsource Detections. *Frontiers in Earth Science*, 8, July 2020. doi: 10.3389/feart.2020.00272.

Bossu, R. and Mazet-Roux, G. An operational authoritative location scheme. *New Manual of Seismological Observatory Practice 2 (NMSOP2)*, 2012. doi: 10.2312/GFZ.NMSOP-2_IS_8.5.

Bossu, R., Mazet-Roux, G., Douet, V., Rives, S., Marin, S., and Aupetit, M. Internet Users as Seismic Sensors for Improved Earthquake Response. *Eos, Transactions American Geophysical Union*, 89(25):225–226, June 2008. doi: 10.1029/2008eo250001.

Bossu, R., Gilles, S., Mazet-Roux, G., Roussel, F., Frobert, L., and Kamb, L. Flash sourcing, or rapid detection and characterization of earthquake effects through website traffic analysis. *Annals of Geophysics*, 54(6), Jan. 2012. doi: 10.4401/ag-5265.

Bossu, R., Lefebvre, S., Cansi, Y., and Mazet-Roux, G. Characterization of the 2011 Mineral, Virginia, Earthquake Effects and Epicenter from Website Traffic Analysis. *Seismological Research Letters*, 85(1):91–97, Jan. 2014. doi: 10.1785/0220130106.

Bossu, R., Steed, R., Mazet-Roux, G., Roussel, F., Etivant, C., Frobert, L., and Godey, S. *The Key Role of Eyewitnesses in Rapid Impact Assessment of Global Earthquakes*, page 601–618. Springer International Publishing, Sept. 2015. doi: 10.1007/978-3-319-21753-6_25.

Bossu, R., Landès, M., Roussel, F., and Steed, R. Felt Reports for Rapid Mapping of Global Earthquake Damage: The Doughnut Effect? *Seismological Research Letters*, 89(1):138–144, Dec. 2017. doi: 10.1785/0220170129.

Bossu, R., Roussel, F., Fallou, L., Landès, M., Steed, R., Mazet-Roux, G., Dupont, A., Frobert, L., and Petersen, L. LastQuake: From rapid information to global seismic risk reduction. *International Journal of Disaster Risk Reduction*, 28:32–42, June 2018a. doi: 10.1016/j.ijdrr.2018.02.024.

Bossu, R., Steed, R., Roussel, F., Landès, M., Fuenzalida, A., Matrullo, E., Dupont, A., Roch, J., and Fallou, L. App Earthquake Detection and Automatic Mapping of Felt Area. *Seismological Research Letters*, 90(1):305–312, Nov. 2018b. doi: 10.1785/0220180185.

Bossu, R., Finazzi, F., Steed, R., Fallou, L., and Bondár, I. "Shaking in 5 Seconds!"—Performance and User Appreciation Assessment of the Earthquake Network Smartphone-Based Public Earthquake Early Warning System. *Seismological Research Letters*, 93(1):137–148, Sept. 2021. doi: 10.1785/0220210180.

Bossu, R., Corradini, M., Cheny, J.-M., and Fallou, L. A social bot in support of crisis communication: 10-years of @LastQuake experience on Twitter. *Frontiers in Communication*, 8, Mar. 2023. doi: 10.3389/fcomm.2023.992654.

Böse, M., Heaton, T. H., and Hauksson, E. Real-time Finite Fault Rupture Detector (FinDer) for large earthquakes: Finite Fault Rupture Detector (FinDer). *Geophysical Journal International*, 191(2):803–812, Sept. 2012. doi: 10.1111/j.1365-246x.2012.05657.x.

Böse, M., Julien-Laferrière, S., Bossu, R., and Massin, F. Near Real-Time Earthquake Line-Source Models Derived from Felt Reports. *Seismological Research Letters*, 92(3):1961–1978, Jan. 2021. doi: 10.1785/0220200244.

Cocco, M., Freda, C., Atakan, K., Bailo, D., Saleh-Contell, K., Lange, O., and Michalek, J. The EPOS Research Infrastructure: a federated approach to integrate solid Earth science data and services. *Annals of Geophysics*, 65(2):DM208, Apr. 2022. doi: 10.4401/ag-8756.

Contreras, D., Wilkinson, S., Aktas, Y. D., Fallou, L., Bossu, R., and Landès, M. Intensity-Based Sentiment and Topic Analysis. The Case of the 2020 Aegean Earthquake. *Frontiers in Built Environment*, 8, Mar. 2022. doi: 10.3389/fbuil.2022.839770.

Earle, P. S., Bowden, D., and Guy, M. Twitter earthquake detection: earthquake monitoring in a social world. *Annals of Geophysics*, 54(6), Jan. 2012. doi: 10.4401/ag-5364.

Fallou, L., Corradini, M., Bossu, R., and Cheny, J.-M. Preventing and debunking earthquake misinformation: Insights into EMSC's practices. *Frontiers in Communication*, 7, Dec. 2022. doi: 10.3389/fcomm.2022.993510.

Finazzi, F. The Earthquake Network Project: Toward a Crowdsourced Smartphone-Based Earthquake Early Warning System. *Bulletin of the Seismological Society of America*, 106(3): 1088–1099, May 2016. doi: 10.1785/0120150354.

Gailler, A., Hébert, H., Loevenbruck, A., and Hernandez, B. Simulation systems for tsunami wave propagation forecasting within the French tsunami warning center. *Natural Hazards and Earth System Sciences*, 13(10):2465–2482, Oct. 2013. doi: 10.5194/nhess-13-2465-2013.

Giardini, D., Van Eck, T., Bossu, R., and Wiemer, S. Networking Research Infrastructures for Earthquake Seismology in Europe. *Eos, Transactions American Geophysical Union*, 89(24):219–219, June 2008. doi: 10.1029/2008eo240003.

Godey, S., Bossu, R., Guilbert, J., and Mazet-Roux, G. The Euro-Mediterranean Bulletin: A Comprehensive Seismological Bulletin at Regional Scale. *Seismological Research Letters*, 77(4): 460–474, July 2006. doi: 10.1785/gssrl.77.4.460.

Godey, S., Bossu, R., and Guilbert, J. Improving the Mediterranean seismicity picture thanks to international collaborations. *Physics and Chemistry of the Earth, Parts A/B/C*, 63:3–11, Jan. 2013. doi: 10.1016/j.pce.2013.04.012.

Grunthal, G. European Macroseismic Scale 1998 (EMS-98). *Cahiers du Centre Européen de Géodynamique et de Séismologie*, page 99, 1998. doi: 10.2312/EMS-98.full.en.

Guérin-Marthe, S., Gehl, P., Negulescu, C., Auclair, S., and Fayjaloun, R. Rapid earthquake response: The state-of-the art and recommendations with a focus on European systems. *International Journal of Disaster Risk Reduction*, 52:101958, Jan. 2021. doi: 10.1016/j.ijdrr.2020.101958.

Haslinger, F., Basili, R., Bossu, R., Cauzzi, C., Cotton, F., Crowley, H., Custodio, S., Danciu, L., Locati, M., Michelini, A., Molinari, I., Ottemöller, L., and Parolai, S. Coordinated and Interoperable Seismological Data and Product Services in Europe: the EPOS Thematic Core Service for Seismology. *Annals of Geophysics*, 65 (2):DM213, Apr. 2022. doi: 10.4401/ag-8767.

Hayes, G. P., Earle, P. S., Benz, H. M., Wald, D. J., and Briggs, R. W. 88 Hours: The U.S. Geological Survey National Earthquake Information Center Response to the 11 March 2011 Mw 9.0 Tohoku Earthquake. *Seismological Research Letters*, 82(4):481–493, July 2011. doi: 10.1785/gssrl.82.4.481.

Hough, S. E., Martin, S. S., Gahalaut, V., Joshi, A., Landes, M., and Bossu, R. A comparison of observed and predicted ground motions from the 2015 MW7.8 Gorkha, Nepal, earthquake. *Natural Hazards*, 84(3):1661–1684, Aug. 2016. doi: 10.1007/s11069-016-2505-8.

Hutko, A. R., Bahavar, M., Trabant, C., Weekly, R. T., Fossen, M. V., and Ahern, T. Data Products at the IRIS-DMC: Growth and Usage. *Seismological Research Letters*, 88(3):892–903, Apr. 2017. doi: 10.1785/0220160190.

Julien-Laferrière, S. Earthquake Qualitative Impact Assessment: Performance evaluation, 2019. http://www.sera-eu.org/export/sites/sera/home/.galleries/pdfs/Performance_analysis_EQIA.pdf. SERA-JRA6 Report (last accessed March 2022).

Kouskouna, V., Ganas, A., Kleanthi, M., Kassaras, I., Sakellariou, N., Sakkas, G., Valkaniotis, S., Manousou, E., Bozionelos, G., Tsironi, V., Karamitros, I., Tavoularis, N., Papaioannou, C., and Bossu, R. Evaluation of macroseismic intensity, strong ground motion pattern and fault model of the 19 July 2019 Mw5.1 earthquake west of Athens. *Journal of Seismology*, 25(3):747–769, Apr. 2021. doi: 10.1007/s10950-021-09990-3.

Lilienkamp, H., Bossu, R., Cotton, F., Finazzi, F., Landès, M., Weatherill, G., and von Specht, S. Utilization of Crowdsourced Felt Reports to Distinguish High-Impact from Low-Impact Earthquakes Globally within Minutes of an Event. *The Seismic Record*, 3(1): 29–36, Jan. 2023. doi: 10.1785/0320220039.

Locati, M., Rovida, A., Albini, P., and Stucchi, M. The AHEAD Portal: A Gateway to European Historical Earthquake Data. *Seismological Research Letters*, 85(3):727–734, May 2014. doi: 10.1785/0220130113.

Masse, R. and Needham, R. NEIC-The national earthquake information center. *Earthquakes Volcanoes (USGS*, 21(1):4–44, 1989.

Michelini, A., Iley, G., Necmioğlu, ⊠ Wotawa, G., Arnold-Arias, D., and Forlenza, G. ARISTOTLE (All Risk Integrated System TOwards The hoListic Early-warning)-European Natural Hazard Scientific Partnership. *EGU General Assembly Conference Abstracts*, 2020. doi: 10.5194/egusphere-egu2020-20016.

Mueller, S. *The European-Mediterranean Seismological Centre (E.M.S.C)*. Strasbourg, 1980.

Ofli, F., Qazi, U., Imran, M., Roch, J., Pennington, C., Banks, V., and Bossu, R. *A Real-Time System for Detecting Landslide Reports on Social Media Using Artificial Intelligence*, page 49–65. Springer International Publishing, 2022. doi: 10.1007/978-3-031-09917-5_4.

Pennington, C., Bossu, R., Ofli, F., Imran, M., Qazi, U. W., Roch, J., and Banks, V. J. A Near-Real-Time Global Landslide Incident Reporting Tool Demonstrator Using Social Media and Artificial Intelligence. *SSRN Electronic Journal*, 2022. doi: 10.2139/ssrn.4075922.

Quinteros, J., Strollo, A., Evans, P. L., Hanka, W., Heinloo, A., Hemmleb, S., Hillmann, L., Jaeckel, K.-H., Kind, R., Saul, J., Zieke, T., and Tilmann, F. The GEOFON Program in 2020. *Seismological Research Letters*, 92(3):1610–1622, Feb. 2021. doi: 10.1785/0220200415.

Quitoriano, V. and Wald, D. J. USGS “Did You Feel It?”—Science and Lessons From 20 Years of Citizen Science-Based Macroseismology. *Frontiers in Earth Science*, 8, May 2020. doi: 10.3389/feart.2020.00120.

Quitoriano, V. and Wald, D. J. Combining Macroseismic Intensity Felt Reports Using ShakeMap. *Proceedings of the 3rd European Conference on earthquake engineering and seismology (3ECEES), URL last accessed April 2023*, 2022. https://3ecees.ro/wp551content/uploads/2022/08/Proceedings_3ECEES_2022.pdf.

Rothé, J.-P. Fifty years of history of the international association of Seismology (1901-1951). *Bulletin of the Seismological Society of America*, 71(3):905–923, June 1981. doi: 10.1785/bssa0710030905.

Roudil, P., Schindele, F., Duperray, P., Gailler, A., Hebert, H., Loevenbruck, A., Gutierrez, E., and Damicis, A. The French Tsunami warning center for the Mediterranean and North-East Atlantic: CENALT. In *EGU General Assembly Conference Abstracts*, pages EGU2013–9636, 2013.

Smith, S. W. IRIS—A university consortium for seismology. *Reviews of Geophysics*, 25(6):1203–1207, July 1987. doi: 10.1029/rg025i006p01203.

Steed, R. J., Fuenzalida, A., Bossu, R., Bondár, I., Heinloo, A., Dupont, A., Saul, J., and Strollo, A. Crowdsourcing triggers rapid, reliable earthquake locations. *Science Advances*, 5(4), Apr. 2019. doi: 10.1126/sciadv.aau9824.

Strollo, A. EIDA: The European Integrated Data Archive and Service Infrastructure within ORFEUS. *Seismological Research Letters*, 92(3):1788–1795, Mar. 2021. doi: 10.1785/0220200413.

Van Noten, K., Lecocq, T., Sira, C., Hinzen, K.-G., and Camelbeeck, T. Path and site effects deduced from transfrontier internet macroseismic data of two recent M4 earthquakes in NW Europe. Nov. 2016. doi: 10.5194/se-2016-150.

Wald, D. J., Quitoriano, V., Dengler, L. A., and Dewey, J. W. Utilization of the Internet for Rapid Community Intensity Maps. *Seismological Research Letters*, 70(6):680–697, Nov. 1999. doi: 10.1785/gssrl.70.6.680.

Weber, B., Becker, J., Hanka, W., Heinloo, A., Hoffmann, M., Kraft, T., and Thoms, H. SeisComP3 — Automatic and interactive real time data processing. *EGU, Geophysical Research Abstracts*, 9: 219, 2007.

REPORT
doi:10.26443/seismica.v3i1.729

SEISMICA

The First Network of Ocean Bottom Seismometers in the Red Sea to Investigate the Zabargad Fracture Zone

Laura Parisi * [1], **Nico Augustin** [2], **Daniele Trippanera** [3], **Henning Kirk** [4], **Anke Dannowski** [2,5], **Rémi Matrau** [1], **Margherita Fittipaldi** [1], **Adriano Nobile** [1], **Olaf Zielke** [1], **Eduardo Valero Cano** [1], **Guus Hoogewerf** [1], **Theodoros Aspiotis** [1], **Sofia Manzo-Vega** [1], **Armando Espindola Carmona** [1], **Alejandra Barreto** [1], **Marlin Juchem** [1], **Cahli Suhendi** [1], **Mechita C. Schmidt-Aursch** [4], **P. Martin Mai** [1], **Sigurjón Jónsson** [1]

[1]PSE Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia, [2]GEOMAR Helmholtz Centre for Ocean Research, Kiel, Germany, [3]Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy, [4]Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, Bremerhaven, Germany, [5]K.U.M.-Kiel GmbH, Kiel, Germany

Author contributions: *Conceptualization*: SJ, LP, DT. *Investigation*: LP, HK, AD, RM, MF, AN, OZ, EVC, GH, AEC, APB, MJ, NA, TA, CS, SMV. *Formal Analysis*: LP. *Visualization*: LP, RM. *Writing - original draft*: LP. *Writing - Review & Editing*: LP, NA, DT, AD, AN, EVC, AEC, MSA, PMM, SJ. *Resources*: MSA, PMM, SJ. *Funding acquisition*: LP, DT, SJ. *Project administration*: LP, SJ.

Abstract The slow-spreading Red Sea rift has been the focus of geophysical investigations in the recent past to study the extension of the oceanic crust, the thickness of the sedimentary cover, and the formation of transform faults. Despite these efforts, local seismology datasets remain scarce, limiting their potential contribution to understanding the tectonic evolution of the Red Sea. The Zabargad Fracture Zone, situated in the Northern Red Sea, offsets the rift axis to the East, making it an important tectonic element to better understand the Red Sea rift's formation. To fill the gap of missing seismological observations, we deployed the first passive seismic network in the Red Sea, specifically within the Zabargad Fracture Zone. This network comprised a total of 14 ocean-bottom seismometers (OBS) and four portable onshore broadband seismic stations, positioned on islands and along the Saudi Arabian coast. Our noise analyses revealed that short-period noise (less than 0.2 s) in this region is more pronounced than in many other areas sampled by OBSs, possibly due to intense ship traffic. Within the microseismic noise range, we identified strong contributions from local atmospheric and oceanic sources of noise, which in combination with site effects generated a second peak around 0.2-1 s. At long periods, waveforms may be used for regional and global studies of earthquakes larger than magnitude $M_w \approx 6.7$, and potentially smaller events for an OBS sub-dataset. Finally, we detected a local earthquake with a magnitude $M_w \approx 3.4$, which could have a volcanic or hydrothermal origin.

Production Editor:
Gareth Funning
Handling Editor:
Stephen Hicks
Copy & Layout Editor:
Théa Ragon

Received:
May 1, 2023
Accepted:
March 28, 2024
Published:
May 1, 2024

1 Introduction

The Red Sea is a slow to ultra-slow spreading ridge with an age of less than 14 million years (e.g., Augustin et al., 2021; Delaunay et al., 2023), formed after the break-up of Arabia from Nubia. While many tectonic models of the Red Sea have limited the extent of oceanic spreading to the southern and central Red Sea (e.g., Coleman and McGuire, 1988; Almalki et al., 2015), increasing amount of evidence is pointing to mid-ocean spreading along its entire length (Augustin et al., 2021; Delaunay et al., 2023). An offset of up to ~100 km marks the transition between the northern and central Red Sea. This is usually referred to as the Zabargad Fracture Zone (ZFZ) and extends from Zabargad Island in the South to the southern limit Mabahiss Deep in the North (Figure 1). While there is agreement on the presence of the rift axis in the Mabahiss Deep and Mabahiss Mons (an active volcano located north of the deep, see Figure 1, e.g., Augustin et al., 2021; Delaunay et al., 2023; Fittipaldi et al., 2024), the limits and geological structures of the ZFZ are unclear, mostly because the northern and central Red Sea seafloor is covered by thick sediments (mainly evaporites and Plio-Pleistocene sediments) with the basement exposed at only a few locations. Determining the structure of the ZFZ as one or multiple transform faults, or even as a set of non-transform offsets (NTOs), has important implications for the maximum earthquake size in the ZFZ and thus for seismic and tsunami hazard assessments of coastal communities in this part of the Red Sea.

Accurate earthquake locations are critical for resolving the structure of the ZFZ. However, the existing earthquake catalog from the Saudi Geological Survey exhibits two diffuse clusters of seismicity that are more than 50 km in diameter (Figure 1), with event locations far from being sufficient to image the ZFZ fault system in detail. Moreover, not much is known about large earthquakes in the ZFZ. Only four earthquakes of magnitude larger than 4.5 were instrumentally recorded in the area before the Saudi and Egyptian seismic networks were established and since then only one additional earthquake was large enough (M_w 4.9 in 2015) for focal mechanism determination, showing normal faulting (Figure 1). Strike-slip tectonic movements are also expected in the ZFZ and have been suggested from

*Corresponding author: laura.parisi@kaust.edu.sa

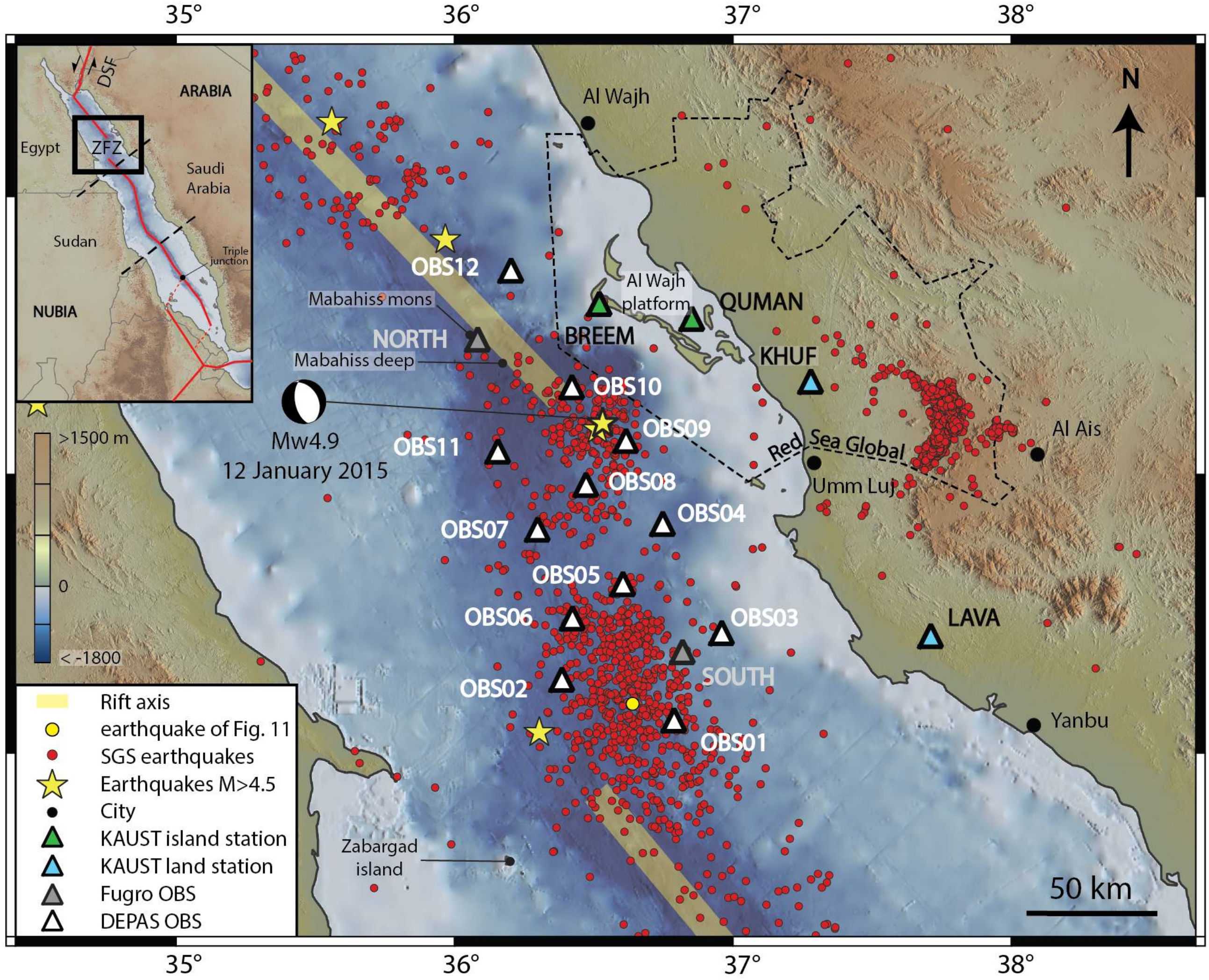


Figure 1 Map of the OBS deployment in the northern Red Sea. The earthquake locations are from the Saudi Geological Survey for 2011-2016, locations of magnitude larger than 4.5 events from IRIS (www.iris.edu, last accessed 5 April 2023), and the focal mechanism of the 12 January 2015, earthquake from the CMT catalog (www.globalcmt.org, last accessed 5 April 2023). Bathymetric data are from GEBCO (www.gebco.net, last accessed 5 April 2023). ZFZ: Zabargad Fracture Zone. DSF: Dead Sea Fault. Dashed lines in the inset separate the northern, central, and southern Red Sea.

tectonic mapping on Zabargad island (Marshak et al., 1992). Furthermore, historical catalogs (El-Isa, 2015; Rehman et al., 2017), while limited, highlight the seismogenic potential of the ZFZ by including reports on two earthquakes of magnitude that could have been as large as M_w 6. Moderate-to-large earthquakes within the ZFZ would threaten neighboring coastal communities on both sides of the Red Sea (Figure 1), in particular on the more populated Saudi coast, e.g., the city of Yanbu with its large petrochemical facilities, the town of Umm Lujj, and Red Sea Global, a major tourist destination under development mostly within the Al Wajh lagoon/platform (Figure 1). The threat may come from the shaking of the weakly consolidated terrains east of the Al-Wajh lagoon and from tsunami waves that could hit the westernmost islands of the same lagoon.

To improve knowledge and understanding of the structure of the ZFZ, we installed the first network of Ocean Bottom Seismometers (OBS) in the Red Sea. We deployed 14 OBSs and four land stations that cover the latitude range 23.5^o-25.5^{o}N (see Figure 1) to establish the temporary ZAFRAN seismic network surrounding the ZFZ to collect broadband seismic data for approximately one year. The primary objective of the deployment is to detect and locate earthquakes to map active faults within the ZFZ. We also plan to construct seismic velocity models of the crust and upper mantle structure to estimate the extent of the oceanic crust and the thickness of the evaporite cover. We will achieve this by utilizing body waves from local and teleseismic events, as well as ambient noise cross-correlations. The results should provide valuable insights into the seismic potential of the ZFZ and, in the broader context, its role in the geological evolution of the Red Sea. Furthermore, the OBS data may help in identifying and characterizing potential active volcanic and hydrothermal sources.

While OBSs enable the collection of data in previously unexplored regions, interpreting and removing seismic noise recorded by OBSs is more challenging compared to noise recorded inland. This complexity arises because sensors deployed inland benefit from insulation against temperature variations, air currents, and wildlife, due to installations in vaults, direct instrument burial, and other protective measures. In contrast, OBSs are often placed directly on the seafloor without specific protection. The exposure of OBSs to the marine elements may introduce noise in the same frequency band as signals used in seismological investigations (ranging from 40 Hz to 100 s and beyond).

Both OBSs and land seismic stations record the ambient seismic noise that occurs in different frequency bands, corresponding to different noise sources. Seismic noise with wave periods exceeding 1 second generally results from the intricate interplay between the atmosphere, the ocean, and the Earth. The microseismic noise range, commonly defined from 2 to 4 seconds up to 20 seconds (e.g., Stutzmann et al., 2000; Gualtieri et al., 2013), encompasses the secondary microseismic peak, often observed around 7 seconds, and sometimes observed split with an additional peak around 1-2 s (e.g., Parisi et al., 2020). Long-period ambient noise (periods exceeding 10-20 seconds), which typically includes the primary microseismic peak occurring between 10 and 20 seconds, is generated through the interaction of ocean waves with shorelines (e.g., Ardhuin, 2018). Periods exceeding 30 seconds are frequently influenced by infragravity waves originating in coastal regions, capable of propagating back to the open ocean (Ardhuin et al., 2014). Detecting and mitigating this noise from OBS data usually involves analyzing the coherence between pressure and vertical seismic signals (e.g., Janiszewski et al., 2019). Additionally, sea bottom currents, circulating the OBS elements, can introduce noise within this period range, predominantly affecting the horizontal components. In cases of imperfect sensor leveling, this noise may impact the vertical component as well (Crawford and Webb, 2000). The detection and removal of this noise can be achieved by examining the coherence, if present, between the vertical and horizontal components (Crawford and Webb, 2000).

The quality of short-period signals (T<1 s), on the other hand, is less dependent on the sensor performance, but it is crucial for the investigation of local earthquakes and volcanic and hydrothermal activities. Nevertheless, OBS recordings have frequently reported additional sources of seismic signals at short periods, including those generated by ships and marine mammals (e.g., Wilcock, 2012; Trabattoni et al., 2023). Moreover, signals within these short periods may be susceptible to corruption by noise, often generated by sea bottom currents interfering with protruding elements of the OBSs, such as the antenna, beacon, and flags (e.g., Stähler et al., 2018; Essing et al., 2021a).

In this article, we present the deployment of the OBS network in the ZFZ, show examples of the collected data, and highlight notable signals recorded. Furthermore, we provide recommendations for utilizing the dataset, drawing from our data quality assessment and analysis. Lastly, our contribution extends to enhancing the comprehension of the splitting of the secondary microseism peak in the microseismic noise band.

2 The ZAFRAN seismic network

We operated the ZAFRAN seismic network from September 2021 to January 2023 with most of the instruments collecting data from November 2021 to November 2022. The network included 14 broadband OBSs and four onshore portable seismic stations, covering the northern Red Sea in the latitude range 24.0^o-25.8^{o}N and from longitude of 36.5^{o}E to the western coast of Saudi Arabia (Figure 1). The OBS interstation distances ranged from 17 to 42 km whereas the onshore stations were more widely spaced (33-110 km), because they were primarily installed to cross-validate the OBSs waveforms. The station coordinate information can be found in Suppl. Table 1.

2.1 Offshore deployment

The offshore part of the network consisted of 12 Lobster OBSs (stations codes from OBS01 to OBS12) from the DEPAS pool (Alfred-Wegener-Institut Helmholtz-Zentrum für Polar-und Meeresforschung et al., 2017) (Figure 2a and 3), deployed at water depths between 740 m and 1700 m (depths are listed in Suppl. Table 1) and two OBSs designed and deployed by Fugro (station codes NORTH and SOUTH; Figure 2b) at depths of 960 m and 870 m, respectively. The DEPAS OBSs have been used in many OBS deployments around the world (e.g., Geissler et al., 2010; Stähler et al., 2016; Blanck et al., 2020), while the Fugro OBS setup is experimental and has not been tested before. Each DEPAS OBS consisted of a 1.65 x 1.30 m frame equipped with a Güralp CMG-40T-OBS sensor and a SEND MCS data logger hosted in titanium pressure-resistant tubes. The sensor was placed between two floating units, mounted to a metallic plate that sits on an anchor (Figure 3a). To facilitate the import of the OBSs into Saudi Arabia, our setup did not include a hydrophone, in contrast to many OBS deployments. Additional equipment to allow and facilitate the instrument recovery included a floating unit, acoustic releaser, flashlight, radio beacon, flag, and buoy (Alfred-Wegener-Institut Helmholtz-Zentrum für Polar-und Meeresforschung et al., 2017), which is attached to the OBS with about 10 m long and 18 mm thick polypropylene rope (Figure 3b). Given that the rope and buoy have been found to be responsible for high-frequency noise (e.g., Stähler et al., 2018), we deployed six DEPAS OBSs by wrapping the rope and the buoy with a fabric fixed to the releaser and the other six with a free rope and buoy to study the difference in the noise properties. The metallic anchor, allowing each OBS to sink to the seafloor, was locked to the frame through the releaser. The DEPAS OBSs were deployed by free-fall from the research vessels R/V Thuwal (KAUST, Thuwal) and R/V Al Azizi (King Abdulaziz University, Jeddah). Due to the limited deck space on these vessels, the deployment was conducted during two short trips from the KAUST harbor in November 2021.

To recover the DEPAS OBSs, an acoustic release command is sent through the water column to the releaser. The releaser then unlocks the anchor from the OBS, which becomes buoyant enough to float to the sea sur-

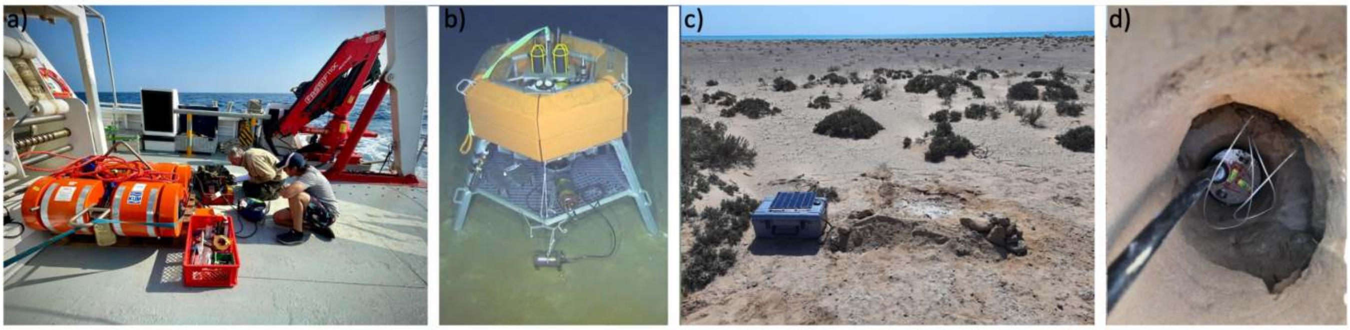


Figure 2 Photos of the seismic equipment used in the ZAFRAN network. a) A DEPAS OBS on the R/V Thuwal. b) Fugro multi-sensor lander with seismometer on the sea bottom. c) Installation of the island station BREEM (see Figure 1; setup identical to onshore stations). The grey box contains the data logger and the batteries. The solar panel and GNSS antenna are placed on the box. The seismometer is buried (not visible) and connected to the logger by the black cable. d) Example (in KHUF) of seismometer installation before filling the hole with sorted sand.

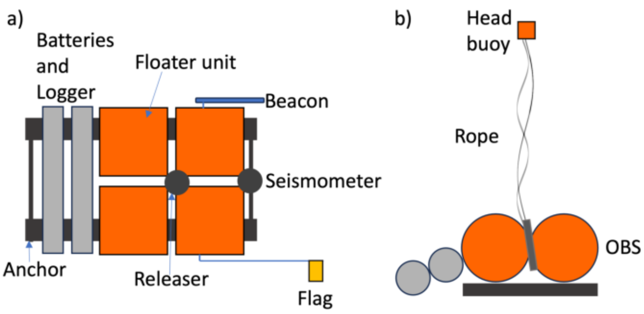


Figure 3 Schematic representation of the DEPAS OBS (Lobster)from a) above and b) from the side. Both sketches are not in scale. A head buoy is attached to the OBS with a free rope; half of the DEPAS OBS had the rope free to strum like in b) and the other half had the rope fixed on the anchor.

face. While 11 of the 12 DEPAS OBSs were successfully recovered during two trips in November 2022, communication with one of the OBSs (OBS04) was not successful, such that an additional trip was required in January 2023 when an automatic release had been scheduled. On 16 January 2023, the OBS04 was recovered without showing any damage, so the reason for the earlier unsuccessful recovery remains unknown. Skew values (difference between the time of the data logger and the instantaneous GNSS time) were measured for all DEPAS OBSs, except for OBS04, and are available in Suppl. Table 1.

Partially overlapping in time with our DEPAS OBS deployment, Fugro conducted an experimental deployment of two multi-sensor deep landers that include OBSs (Figure 2b). The Fugro OBSs were deployed in September 2021 from S/V Kobi Ruegg and visited in February and July 2022 with OSS Handin Tide. During the visits, the landers were fully recovered and redeployed after data collection and instrument maintenance. The OBS setup and deployment protocol for the Fugro OBSs were different from the DEPAS OBSs. After the landers reached the seafloor, an ROV was used to place a Nanometric Trillium Compact OBS 120s seismometer at the seafloor. It was enclosed on a light frame with feet to couple with the seafloor sediments. The seismometer was in an aluminum casing that weighs 2.9 kg in water. The data-logger Nanometric Pegasus OBS and batteries were on the lander and connected to the sensor with a cable.

During the visit in February 2022, the SOUTH data logger was found to have a minor leakage, and no data were recovered due to a damaged cable. The OBS from the lander NORTH was recovered and deployed on the lander SOUTH. Data from SOUTH were then finally recovered in July 2022. The skew values are not available.

Skew values for the Lobster OBSs range from 0.01 to 1.3 s, with a median of 0.37 s (Suppl. Table 1). The highest skew value was found for OBS08 that, together with the issues described in section 2.3, may indicate a possible general malfunction of the instrument. When excluding OBS08, the median (mean) skew value is 0.33 s (0.34 s). Although the skew values for OBS04 and NORTH are missing, these can be recovered using ambient noise cross-correlations between onshore and offshore stations (e.g., Naranjo et al., 2024).

We determined the orientation of the horizontal components of the off-shore seismometers using two distinct, data-type-based methods implemented in the open-source Python package (OrientPy; https://github.com/nfsi-canada/OrientPy). The first method is based on minimizing the P- and PP-wave energy on the transverse component (Braunmiller et al., 2020) while the second method is based on the arrival angle of minor- and major-arc intermediate-period surface-waves of teleseismic earthquakes by using modern global dispersion maps (Doran and Laske, 2017). For each station, we select the orientation according to the method with the smaller uncertainty (Suppl. Table 1). The method based on the polarization of surface waves yields smaller uncertainties for almost all the stations, except for OBS08 and OBS10. For benchmark and completeness, we also calculate the orientation of the onshore stations that resulted to be always lower than 10^o. The final median of uncertainties is 12.2^o.

2.2 Onshore deployment

We complemented the OBS network by installing four onshore stations. Two of them were located on the

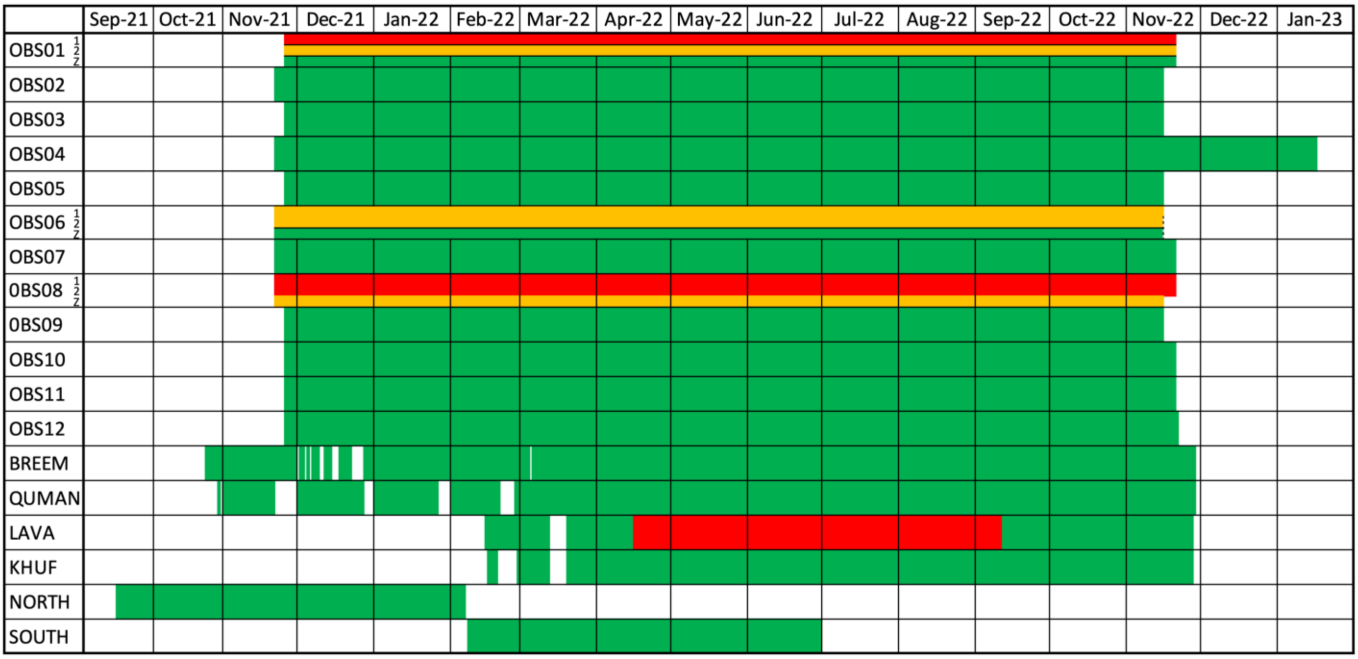


Figure 4 Data availability of the ZAFRAN dataset from September 2021 to January 2023. Green corresponds to waveforms available in three components. Yellow corresponds to data needing further preprocessing before being used. Red represents station or component failure. Note that there are three lines for each station, for the vertical (Z) and two horizontal (1 and 2) components.

small uninhabited reef islands Quman (3 km wide) and Breem (6 km wide) of the Al Wajh platform (Figure 1 and 2c). We refer to these stations as the "island stations". Station QUMAN is located within the Al Wajh lagoon whereas BREEM is located on the edge of the platform, making it more exposed to open sea environmental conditions.

The other two stations were installed onshore at a distance of 15 km (KHUF) and 25 km (LAVA) from the coast (Figure 1). Selecting locations closer to the coast was not possible because of coastguard permit limitations and lack of solid bedrock. We refer to KHUF and LAVA as the "land stations" and to both the island and land stations as the "onshore stations".

The island stations were equipped with Nanometrics Trillium Compact Horizon sensors and the two land stations with Nanometrics Trillium Compact posthole sensors. Both types of sensors have a flat response of up to 120 s and can be used in direct burial installations. The sensors were buried within a depth of 50 cm in a cylindrical hole that was 2 cm larger in diameter than the sensor (Figure 2d). The bottom of the hole was filled with a thin layer of sorted fine sand to easily level the sensor. The little space remaining between the hole and the sensor was filled with the same sand providing coupling and thermal insulation. The two island sensors were deployed within porous, but hard, coral rocks, and the two land sensors were installed within the Precambrian bedrock.

The onshore stations were equipped with Nanometrics Centaur dataloggers powered by lithium batteries charged by a 30 x 40 cm solar panel (Figure 2c). Sand accumulation on solar panels is a well-known issue, especially in this part of the world, and since a definitive solution has not been found yet, data recording has suffered a few gaps because of power issues. Stations were visited for maintenance, and data were collected every 3-6 months.

2.3 The collected dataset

The ZAFRAN dataset includes about 12 months of data from the DEPAS OBSs, 5 months from the Fugro OBSs, 14 months from the island stations, and 10 months from the land stations (Figure 4). More specifically, we collected data for 358 overlapping days with the 12 DEPAS OBS. OBS04 recorded 57 days more than the other DEPAS OBSs. While all the instruments were equipped with 3-component sensors, the quality of the horizontal component 1 of OBS01 and both horizontal components of OBS08 is poor and cannot be used for seismological investigations.

In addition, the vertical component of OBS08 can only be used for half of the recording days (this issue is discussed further in section 3.1). Even considering these data losses, the recovery rate for the DEPAS OBSs is over 90%.

The data of OBS NORTH complements the ZAFRAN dataset with 140 days that overlap for 73 days with the DEPAS dataset. Data from OBS SOUTH span 144 days and completely overlaps with the DEPAS dataset. Considering the initial plan (see section 2.1), the Fugro OBSs had a recovery rate of 50%. Furthermore, the two island stations recorded about 330 days of data, each with an average recovery rate of 96%, and fully overlapping with the DEPAS dataset. Finally, the two land stations provided 270 (KHUF) and 140 days (LAVA) of data at a recovery rate of 95% (KHUF) and 49% (LAVA).

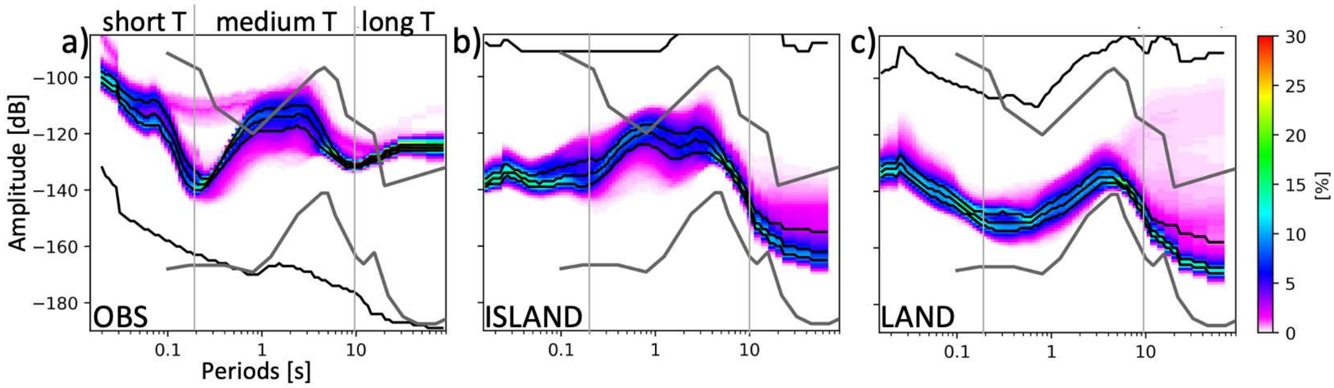


Figure 5 Examples of Probabilistic Power Spectral Densities (PPSD). a) PPSD of the Z component of the station OBS10 calculated for all available data. b) Same as for a) but for the island station BREEM. c) As for a) but for the land station KHUF. Dark gray curves represent the New High Noise Level and the New Low Noise levels (McNamara and Buland, 2004). Black lines correspond to the 25th, 50th (noise level), 75th and 100th. Light gray vertical lines represent the boundaries separating ranges of short, medium, and long periods (T).

3 Noise levels

Noise levels serve as valuable indicators for assessing station performance across different components and for investigating the sources of ambient seismic noise recorded at specific stations. While it is not always straightforward to distinguish the impact of a station's low performance from that of a strong noise source, we discuss the characteristics of noise levels that are primarily associated with instrument type and installation in Section 3.1, and we conduct a preliminary analysis of the noise sources (Section 3.2) to distinguish them from potential issues related to station performances.

To accomplish this, we calculate the Probabilistic Power Spectral Densities (PPSD McNamara and Buland, 2004) as implemented in ObsPy (Beyreuther et al., 2010) for all available data using time windows as small as 450 s and an overlap of 50 %. Examples of PPSD for an OBS, an island, and a land station are shown in Figure 5. We refer to noise level as the median of the PPSD for a given time window (see Figure 5). If not specified, we refer to the entire deployment period of a given instrument.

In our analyses, we divide the overall period range into three segments: short-period ($T < 0.2$ s), medium-period ($0.2 \leq T \leq 10$ s), and long-period ($T > 10$ s). We select the boundaries of 0.2 and 10 seconds due to their alignment with the two predominant notches observed in the PPSD of the ZAFRAN network (see gray vertical lines in the plots of Figure 5). All noise levels are shown in Figure 6 where stations are grouped in classes, depending on the shape of the noise level in the medium-period range. These classes are further discussed in Section 3.2.2.

3.1 Station performances

The two onshore stations, KHUF and LAVA, exhibit the best data quality within the ZAFRAN network, characterized by consistently low noise levels (Figure 6a-c). Their exceptional performance can be attributed to their good subsurface coupling, effective insulation, and remote locations (far from anthropogenic noise sources). However, LAVA displays elevated long-period noise across all components compared to KHUF. This discrepancy may be attributed to thermal insulation limitations, which are also responsible for sensor failures due to high temperatures (see Section 2.2).

The performance of the two island stations, BREEM and QUMAN (Figures 6d-f) is similar to the two land stations because they share the same instruments and style of installation. However, QUMAN exhibits an unusual peak at approximately 4-6 Hz. A visual inspection of waveforms and spectrograms reveals consistent, high-amplitude noise between 10 Hz and 0.8 s, most likely due to construction activities at Red Sea Global. This aspect needs to be considered when using the data for local seismicity studies.

All ZAFRAN stations exhibit a noise level notch between 9 and 11 s (Figure 6). Beyond this period, noise levels consistently increase for nearly all DEPAS OBSs in all components. In contrast, Fugro OBSs maintain low long-period noise in the vertical component, comparable to onshore stations. However, the horizontal components of Fugro OBSs exhibit high noise levels, similar to DEPAS OBSs. These results are in agreement with Stähler et al. (2018) who compared the noise recorded by seismometers deployed inland and offshore to test the Lobster OBSs of the DEPAS pool managed by the Alfred-Wegener Institute. The authors analyzed the noise recorded by the Güralp CMG-40T seismic sensor in vault conditions and the Güralp CMG-40T-OBS at sea. This CMG-40T-OBS is the same sensor but modified to be included in the OBS. They found that the self-noise of the CMG-40T-OBS is higher than the noise produced by the corresponding original model at wave periods larger than 10 s. Their tests also demonstrated that the DEPAS OBSs perform better at longer periods if a Nanometrics Trillium compact seismometer substitutes the CMG-40T-OBS.

Fugro outperforms DEPAS OBSs in the short-period range, likely due to the fewer additional elements on the OBSs that could resonate with marine currents, as discussed in previous studies (Stähler et al., 2016; Essing

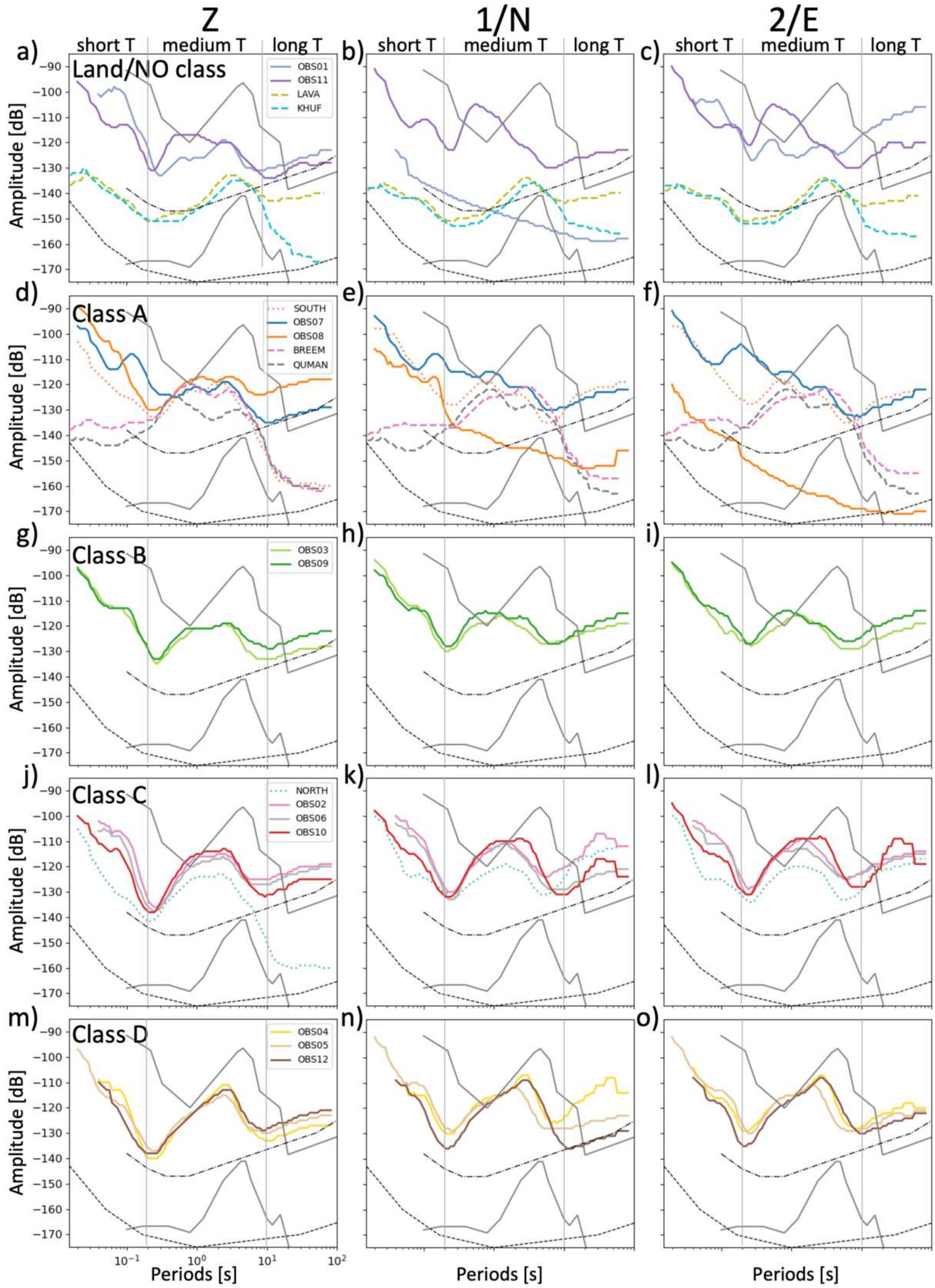


Figure 6 Noise levels (median of PPSD) for the ZAFRAN dataset grouped by similarity in the medium-period range. a) Noise levels for the land stations and two OBS that have unique noise levels for stations of class A and vertical components. b) As in a) but for 1/N components. c) As in A but for Z/E components. d), e) and f) As in a), b) and c) but for the stations in class A. g), h) and i) As in a), b) and c) but for the stations in class B. j), k) and l) As in a), b) and c) but for the stations in class C. m), n) and o) As in a), b) and c) but for the stations in class D. Colored dotted, solid, and dashed lines represent the noise levels of the Fugro OBSs, DEPAS OBSs, and onshore stations, respectively. Noise levels of LAVA do not include days of sensor failure. Gray lines represent the New Low Noise Model and New High Noise Model, respectively (Peterson, 1993). Black lines represent the self-noise of the CMG-40T-OBS (dot-dashed, Stähler et al., 2018) and the Trillium compact (dashed, from the manufacturer).

et al., 2021a; Corela et al., 2022).

However, as already observed by Janiszewski et al. (2022) and valid in our deployment, it is not trivial to separate the effects due to the type of seismometers from the overall OBS setup in the noise level. This also applies to our deployment of two OBS setups with two different sensor types.

An overview of signals recorded by the ZAFRAN network, given in terms of spectrograms calculated from the PPSD (Figure 7), offers further insights into the data quality. The most prominent signal across all stations is the microseisms in the medium-period range (see example in the red "MS" box of Figure 7). Additionally, teleseismic events are visible at several stations in the medium and long-period ranges (see an example within the red box "TL" in March 2022 in Figure 7). Furthermore, local earthquakes are visible in the short-period range (see an example red box "LOC" on 30 June 2022 in Figure 7). These distinct and clear signals serve as evidence of the high-quality nature of the ZAFRAN dataset.

The analysis of noise levels and 1-year spectrograms also highlight sensor failures. The noise levels of the horizontal components 1 of OBS01 and 1 and 2 of OBS08 show that these seismometers' components most of the time did not record properly (Figure 6b and 6e-f). The OBS08 spectrogram (Figure 7) illustrates that the sensor malfunctioned also in the vertical component for approximately 40% of the installation duration. Similarly, the sensor at station LAVA experienced a failure from mid-April to mid-September (Figure 7), likely due to elevated air temperatures (see Section 2.2). Lastly, when examining spectrograms for onshore stations (LAVA, KHUF, BREEM, and QUMAN in Figure 7), we observe intermittent data gaps, which we suspect to be due to sand accumulation on the solar panels. The overall usability of the dataset is summarized in Figure 4.

3.2 Environmental and geological noise sources

Noise levels can also provide insights into environmental (ocean and atmosphere) and geological (subsurface) factors. In this section, we present a preliminary analysis of potential noise sources in the short, medium, and long-period ranges that can be used as a reference for future studies based on the ZAFRAN dataset.

3.2.1 Short-period ranges

Sources of short-period (≤ 0.2 s or ≥ 0.5 Hz) noise can be geological (e.g., local earthquakes, volcanic tremors, etc.) or due to the interaction of sea-bottom currents with the structural components of the OBS (Corela et al., 2022). In addition, noise due to passing ships and marine mammals must be considered.

In general, the short-period noise levels of the ZAFRAN deployment are overall high when compared to PPSDs published from previous experiments in the oceans and lakes (e.g., Stähler et al., 2016, 2018; Hilmo and Wilcock, 2020; Carchedi et al., 2022; Kim et al., 2023; Zhang et al., 2023). For example, ZAFRAN noise levels from the OBSs on the vertical component are in the range between -125 and -110 dB (Figure 6). Most of the noise levels recorded in the Indian Ocean by using the DEPAS OBS (Stähler et al., 2016) and in the South Atlantic (Zhang et al., 2023) are about -130 dB in the same component and frequency. Off the coast of the Pacific Northwest, the noise is between -160 and -150 dB (Hilmo and Wilcock, 2020). The same range of values is found in western Pacific (Kim et al., 2023). In the Malawi Lake, values are about -150 dB (Carchedi et al., 2022). Instead, our short-period noise levels are similar to the noise recorded in shallow water (22 m) of the Baltic Sea (Stähler et al., 2018).

To better understand the origin of the high short-period noise and to identify potential geographical noise patterns within the ZFZ, we plot the noise levels averaged within the short-period range for the vertical and the average of the two horizontal components (Figure 8a and 8e). The geographical noise distribution in the short-period range defines different domains and subdomains. The southern and central offshore domains (labeled "SO" and "CO" in Figures 8a and 8e) reveal higher short-period noise than the northern offshore ("NO") domain for both vertical and horizontal components (Figures 8a and 8e). The limit between these two domains corresponds to the Mabahiss deep (see Figure 1) and the commonly used limit between the central and the northern Red Sea. Finally, for the CO domain, we notice an increase in noise from the island stations offshore for the vertical component. Global maps depicting ship route density (from marinetraffic.com, last access Dec 13, 2023) reveal that the Red Sea is among the most traversed routes, with common shipping routes situated closer to our deployment south of Mabahiss and gradually moving farther north. This fact suggests that ship traffic may contribute to the elevated noise levels and its variation towards the north and offshore.

Another interesting feature of the short-period noise at almost all OBS is a peak at around 0.1 s (10 Hz, Figure 6). Exceptions are OBS05, OBS08 and OBS12. This could be due to a poor coupling between the seismometer and the anchor or between the anchor and the seafloor. However, the reasons for the presence or absence of this peak remain unclear.

Next, we analyze the correlation between noise levels and water depth for the vertical components and for the average of the two horizontal components (ρ^V_{depth} and ρ^H_{depth}, respectively; see Figures 8d and 8h). We group the stations by OBS setup (Fugro and/or DEPAS and loose/tight rope) and we consider only correlation coefficients larger than 0.5. For the short-period noise levels, we find a positive correlation for the horizontal components of the DEPAS OBSs with loose rope (ρ^H_{depth}=0.73) suggesting that the rope may act as a resonant noise-generating element, possibly due to sea-bottom currents. However, given the limited number of stations, this correlation of noise with water depth may not be robust.

Regarding temporal variations of short-period noise levels, all stations exhibit similar patterns across all components. For most stations, the noise remains stable or gradually decreases during the deployment, typically not exceeding a 10 dB change (Figure 9), probably

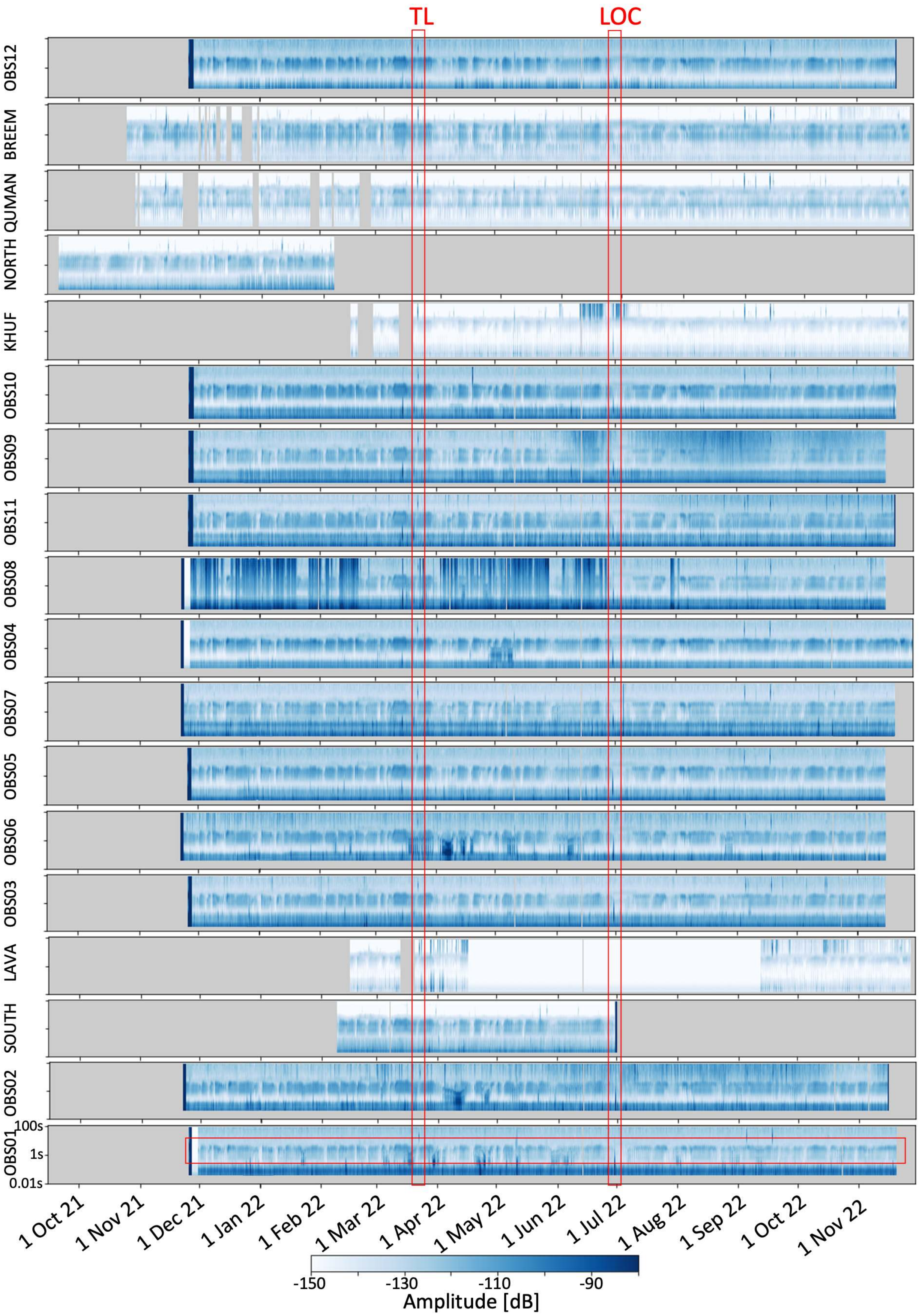


Figure 7 Deployment overview in terms of spectrograms, showing the temporal variations of the PPSD for the vertical component of each station. The vertical scale of each spectrogram is as in the bottom one. Stations are ordered from South (bottom) to North (top).TL: example of teleseismic earthquake. LOC: example of a local earthquake. MS: example of microseismic noise.

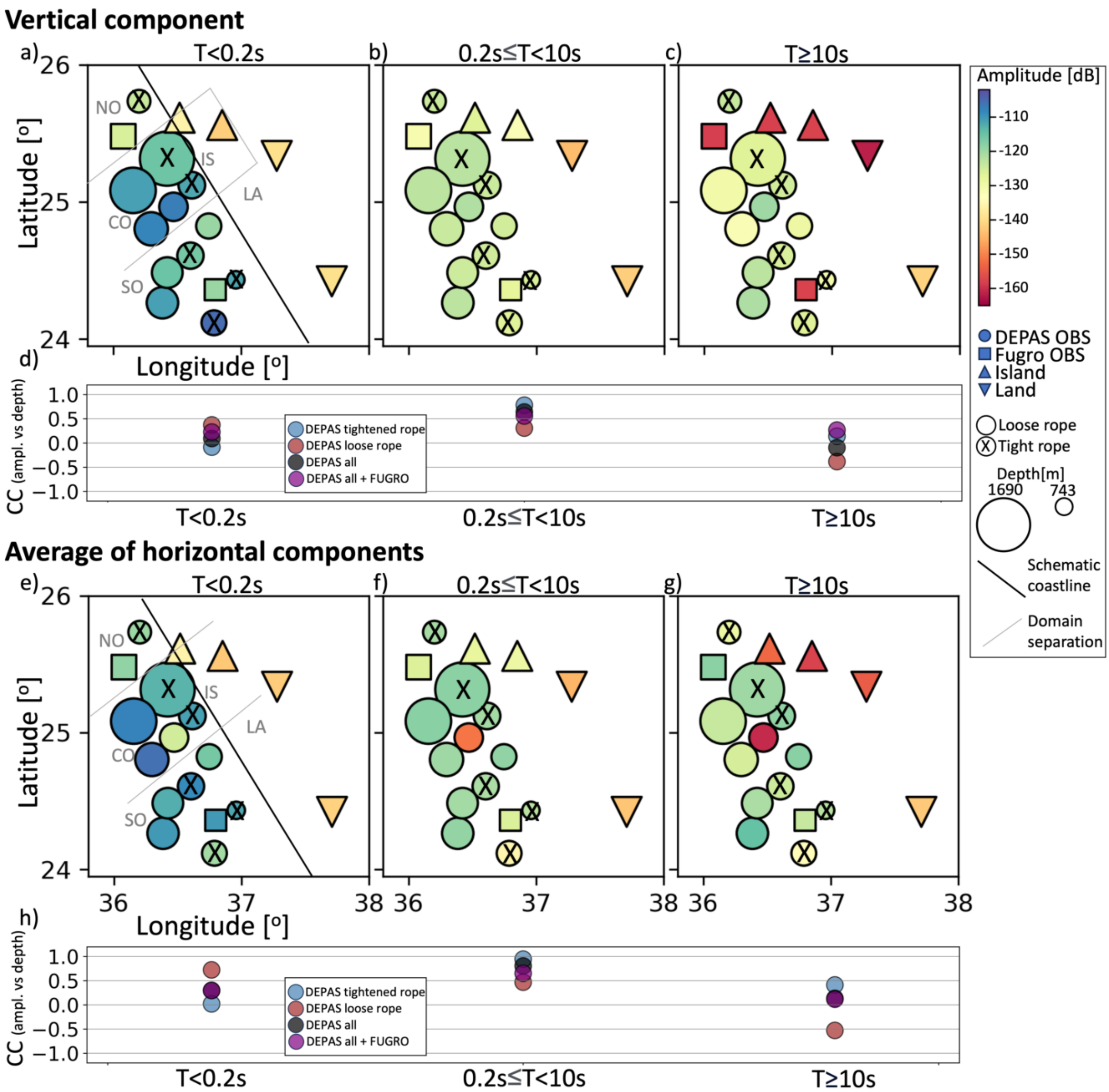


Figure 8 Spatial variations of noise levels. a) Average noise levels for the vertical components at period $T < 2\,s$. NO: northern offshore; CO: central offshore; SO: southern offshore: IS: island/lagoon; LA: land. OBS markers are scaled by water depth. b) and c) as for a) but for periods 0.2 - 10 s and $\geq$10 s, respectively. d) Correlation coefficients between the noise levels on the vertical components and water depth. e), f) and g) as for a), b) and c), respectively, but for the horizontal components. g) As for d) but for the horizontal components.

due to the settling of the instruments on the seafloor becoming more stable. One exception is QUMAN, where short-period noise increases during deployment, likely due to the anthropogenic noise (see section 3.1). Another exception is NORTH, which shows a sharp noise increase of approximately 10 dB from December 2021 to January 2022. Given that NORTH is located on Mabahiss Mons, whose volcanic activity is unknown, it is challenging to determine whether this increase results from changes in volcanic activity, instrument issues, or another unknown phenomenon.

3.2.2 Medium-period range

The secondary microseismic peak is visible on all ZAFRAN stations typically between 1 and 4 s (0.25 - 1 Hz). The island stations and most OBSs also exhibit a second peak at periods between 0.2 and 1 s (1 and 5 Hz, Figure 6). However, this peak is not visible at the land stations (see comparison in Figure 5) and at OBS04, OBS05, and OBS12 (Figure 6m-o). To better understand these differences, we visually classify stations based on the shape of their noise level patterns in the medium-period range, primarily related to the presence or absence of the second peak and the degree of overlap between the two peaks.

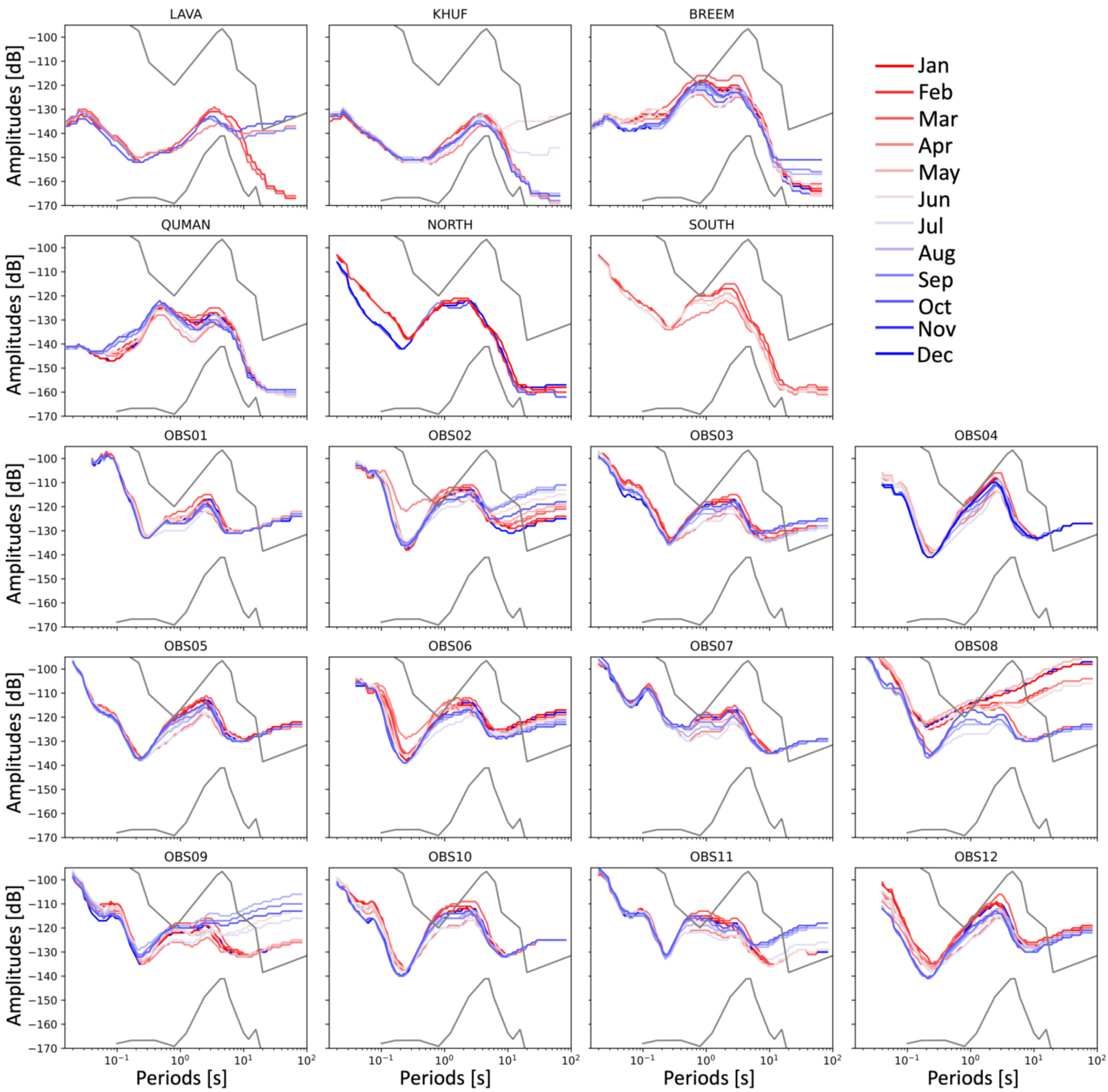


Figure 9 Temporal (monthly) variations of noise levels. Noise levels of the vertical component of all ZAFRAN stations are calculated in monthly time windows. Gray lines represent the New Low Noise Model and New High Noise Model, respectively (Peterson, 1993).

Class A includes stations with two well-separated peaks, such as the island stations QUMAN and BREEM, the Fugro OBS SOUTH, and DEPAS OBS OBS07 and OBS08 (Figure 6d-f). Stations with slightly separated peaks, like OBS03 and OBS09, fall into class B (Figure 6g-i). Class C encompasses stations with noise levels displaying a single large peak rather than two separate peaks (Figure 6j-l), including NORTH, OBS02, OBS06, and OBS10. OBS04, OBS05, and OBS12, which lack a visible second peak, are classified as class D. OBS01 and OBS11 did not fit into any of the predefined classes (Figure 6a-c). Notably, we could not identify any correlations between classes and water depth, geographical location, or instrument type/setup.

To investigate potential correlations between the seismic noise and oceanographic and meteorological phenomena, we use two ERA5 datasets from ECMWF (European Centre for Medium-Range Weather Forecasts, Hersbach et al., 2023, last access 29 August 2023). These ERA5 datasets are reanalyses, consisting of hourly time series and combining models with observational data. Specifically, we examined the 10-meter vertical wind component (10v) with a spatial resolution of 0.25^o, representing northward wind at a height of 10 meters, and the significant wave height (swh), which combines wind waves and swell, also at a resolution of 0.25^o. We use the Pearson coefficient to estimate the level of correlation between the time series of the noise, hourly resampled, and the time series of 10v and swh.

In our analysis, we use two weeks of data (1-14 January 2023 or 1-14 April 2023, depending on the availability) from the DEPAS OBS12, the Fugro OBS NORTH, the two island stations (BREEM and QUMAN), and the land station KHUF and we compare them with 10-meter vertical wind component (10v) and the significant wave height (swh, only used for OBSs data analysis) for the same time windows. Figure 10a shows an example of time-series comparisons for the noise levels of the OBS12 at 0.7 s, 10v and swh at the same location. After calculating the Pearson coefficients for the two parameters (ρ_{10v} and ρ_{swh}) for the full period range, we observe that both curves ρ_{10v} and ρ_{swh} exceed 0.75 in the medium-period range (Figure 10b-c) and are low outside the medium-period range. Also, while the maximum correlation is higher for swh, both correlation curves exhibit a double peak. For both OBSs, the correlation with swh peaked at 1 and 3 s, while the correlation with 10v peaked at 0.5 and 2 s. The periods of the two peaks in the correlation curves align with the periods of the two peaks of the noise levels in the medium-period range observed in most of the ZAFRAN stations. Interestingly, it is worth noting that even though OBS12 belongs to class D (no second peak), ρ_{10v} and ρ_{swh} still exhibit the double peak. For the island stations, ρ_{10v} is slightly lower than that for the OBS. Although the two peaks in the correlation curve are less pronounced, the overall curve shape is similar (Figure 10d-e). For the land station, ρ_{10v} is less than 0.25 but two peaks and the overall shape of the curve are preserved (Figure 10f).

Time-series environmental correlations between noise levels and wind speeds (e.g., Bromirski et al., 2005; Hilmo and Wilcock, 2020) and significant wave height (Zhang et al., 2023; Kim et al., 2023) were already present in literature. Our analysis adds insight on the atmosphere-ocean-Earth interaction clearly showing the frequency domain of influence on the medium-period range noise of these two environmental factors, that for the Red Sea is between 0.2 and 10 s.

Maps in Figures 8b and 8f show small noise variations within the medium-period range with some increase of the noise noticeable going from land stations (LA) to the inner lagoon (easternmost station in IS) to the external lagoon (westernmost station in IS) to offshore.

Regarding the correlation between the noise levels and water depth at the medium-period range, ρ_{depth} is larger than 0.5 for all groups and components considered, except for the DEPAS OBSs with a loose rope. Considering the correlation found between the noise level and the 10v and swh, it is surprising that the noise may increase with water depth. As for the short-period range, these correlations may not be robust enough.

The monthly noise variations at medium-period range across the components are similar for most of the stations, with differences of less than 15 dB between the noisiest and quietest months (Figure 9). During the summer months of the northern hemisphere, typically May, June, and July, we observe a decrease in noise levels (Figure 9). OBS01, OBS02, and OBS06 stand out because noise reduction occurs primarily in July. Interestingly, while the amplitudes vary in time, the presence or absence of the double peak does not (no change of class in time).

3.2.3 Long-period range

To understand the contribution of the sea bottom currents and tilt on the long-period noise, we use the coherence between the vertical and the two horizontal components calculated by using a modified version of the open-source OBS tools code (Janiszewski et al., 2019) that implements the method of Crawford and Webb (2000). Specifically, we calculate daily coherence as a function of the periods for each OBS for the whole duration of the deployment and consider the median coherence (analogously to the noise levels) between the vertical component of each station and the two horizontal components. Finally, we calculate the mean value for periods > 10 s. We observe a very low coherence for both DEPAS OBS (average between stations and component is 0.07) and Fugro OBSs (average is 0.1). While the low coherence agrees with the low noise levels in the vertical components of Fugro OBSs, the high noise in their horizontal components likely results from sea bottom currents. These instruments can automatically level, regardless of initial tilt, preventing horizontal noise from affecting the vertical component. Conversely, the high noise levels in both the vertical and horizontal components of DEPAS OBSs, coupled with low coherence, are probably due to the high self-noise.

We observe that one or both horizontal components of the OBS02 and OBS10 (Figures 6b-c) show a peak between 20 and 50 s. This is a stable feature for the duration of the deployment but for which we currently have no explanation. KHUF, BREEM, and QUMAN are the only stations with a weak primary microseism peak (10-12 s) visible on the vertical components, probably since the high self-noise at periods larger than 10 s is higher than the typical average amplitudes of the primary microseismic peak at the offshore stations.

The observed long-period noise patterns in Figures 8c and 8g suggest that nearby stations tend to have similar noise levels on both vertical and horizontal components. For instance, OBS06 and OBS02, two closely located stations, exhibit the highest noise levels in the vertical component.

Lastly, the most notable feature from the maps in Figures 8c and 8g is the correlation with the instrument types discussed above with the Fugro OBS having systematically lower noise than DEPAS OBS in the horizontal components, and with the onshore stations performing systematically better than offshore stations.

In the long-period range, the only strong correlation between noise and water depth is ρ^H_{depth}=-0.56 for the DEPAS OBSs with the free rope, as already seen for the short-period range (Figure 10f and 10j). The negative correlation between noise and water depth aligns with findings in previous studies (e.g., Janiszewski et al., 2022).

Long-period noise levels remain relatively stable with time for most stations (Figure 9). OBS09 and OBS11 experienced sudden increases of 15-20 dB after the summer period, while OBS06 and OBS12 gradually de-

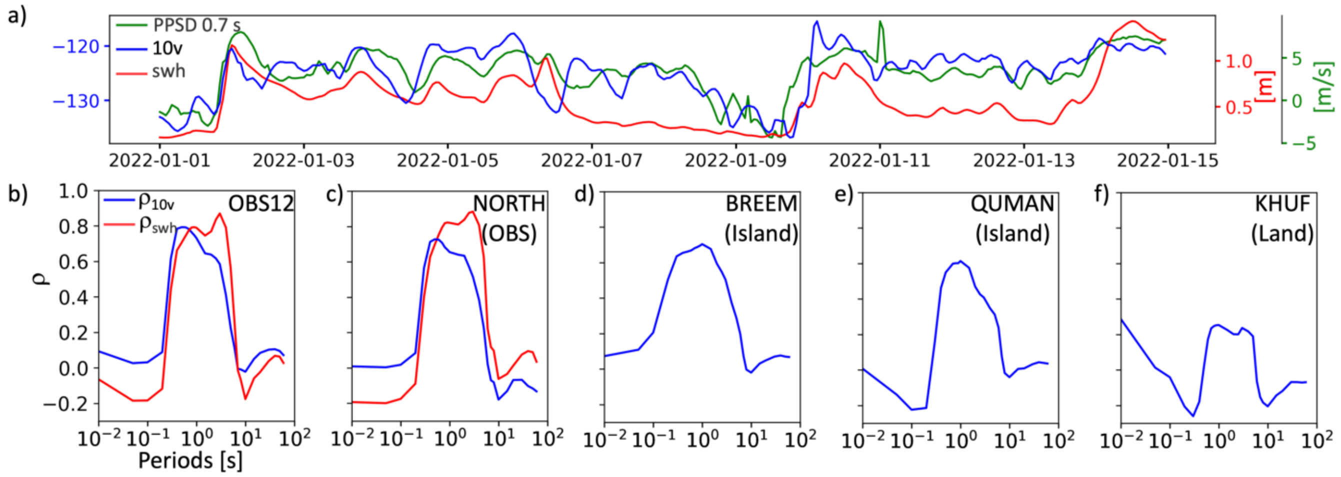


Figure 10 Correlation between environmental data and noise levels. a) Example of comparison of 14-week-long (1-14 January 2022) time series of vertical-component noise levels at the period of 0.7 s for OBS12 against 10-meter vertical wind component anomaly (10v) and significant height of combined wind waves and swell (swh) at the OBS12 location. b) Variation of the Pearson coefficient between noise levels at Fugro OBS NORTH and 10v and swh with the noise level. c) As in b) for DEPAS OBS12. d) As in a) but for island station BREEM (only 10v data available). e) As for d but for island station QUMAN. f) As for d) but for land station KHUF.

creased in noise levels during the deployment, possibly due to instrument settling. OBS02, on the other hand, exhibited a 15-20 dB increase during the summer months, followed by a return to initial values.

4 Notable signals

In this section, we present and briefly discuss some notable seismic signals found in the ZAFRAN dataset that are partly related to the specific setting of the Red Sea and partly due to the deployment and instruments used.

4.1 Long-lasting high-frequency tremor

A notable signal observed within the ZAFRAN network is a recurring, several-hour-long tremor in the 0.025 - 0.25 s period range (4 - 40 Hz). To show this, we plot several one-day spectrograms of the 1/N components for some OBSs and onshore (island and land) stations (Figure 11). We highlight this long-lasting high-frequency tremor by using a dotted-line box on the spectrogram where it is most visible (OBS12). Nonetheless, this signal displays strong amplitudes in the majority of stations, whether offshore or onshore, with the exceptions being QUMAN (the island station), LAVA (the land station), and OBS11. In the spectrograms in Figure 11, this signal starts at around 3 h and it continues for the rest of the day. Although systematic detection has not been conducted as of yet, this type of event is frequently observed in the ZAFRAN datasets, occurring at different times and enduring for several hours. Since the waveforms for these events do not show a sharp onset, their localization is not trivial and is not performed at this stage.

Potential sources of these signals are anthropogenic, such as due to the passage of ships, active seismic surveys, or may originate from various factories (e.g., desalination, refining, cement production) situated along the coast of the study area, particularly in the areas of Al Wajh, Umm Luj, and Yanbu (Hamieh et al., 2022) (see locations in Figure 1). If this is the case, the higher amplitudes observed at the offshore stations might be attributed to more efficient energy propagation over water, possibly facilitated by T-waves traveling through the ocean sound channel (Heleno et al., 2006). Nonetheless, we cannot rule out the possibility of a natural origin (e.g., volcanic tremor from Mabahiss Mons). Heleno et al. (2006) observed tremors with comparable characteristics (duration, frequency, and propagation efficiency) in the Cape Verde islands and proposed active seamounts as potential sources.

4.2 Free rope signature

As described in Section 2.1, half of the DEPAS OBSs were deployed by tightening the rope of the buoy (avoiding free strumming in the water column), and the other half was deployed leaving the rope free to strum on the water column above the OBS (see Suppl. Table 1) and a sketch of the OBSs with free rope in Figure 3b). One-day-long spectrograms of OBSs with the free rope (Figure 11, right column) show two types of noise features that are not visible in the spectrograms of the stations with the tightened rope.

The first type of noise is visible only on the spectrograms of the stations with the free rope is recorded at frequencies above 5 Hz (0.2 s, the most evident example is highlighted in the dashed-dotted line box on the spectrogram of the DEPAS OBS02 in Figure 11). It is characterized by stronger amplitudes at specific frequencies and energy bursts of short duration repeated in time. All DEPAS OBSs with the free rope show this type of noise with some differences in the maximum amplitudes. While this disturbance was investigated by Stähler et al. (2018); Essing et al. (2021b); Corela et al. (2022) and identified between 1 and 10 Hz (0.1 and 1 s) here we

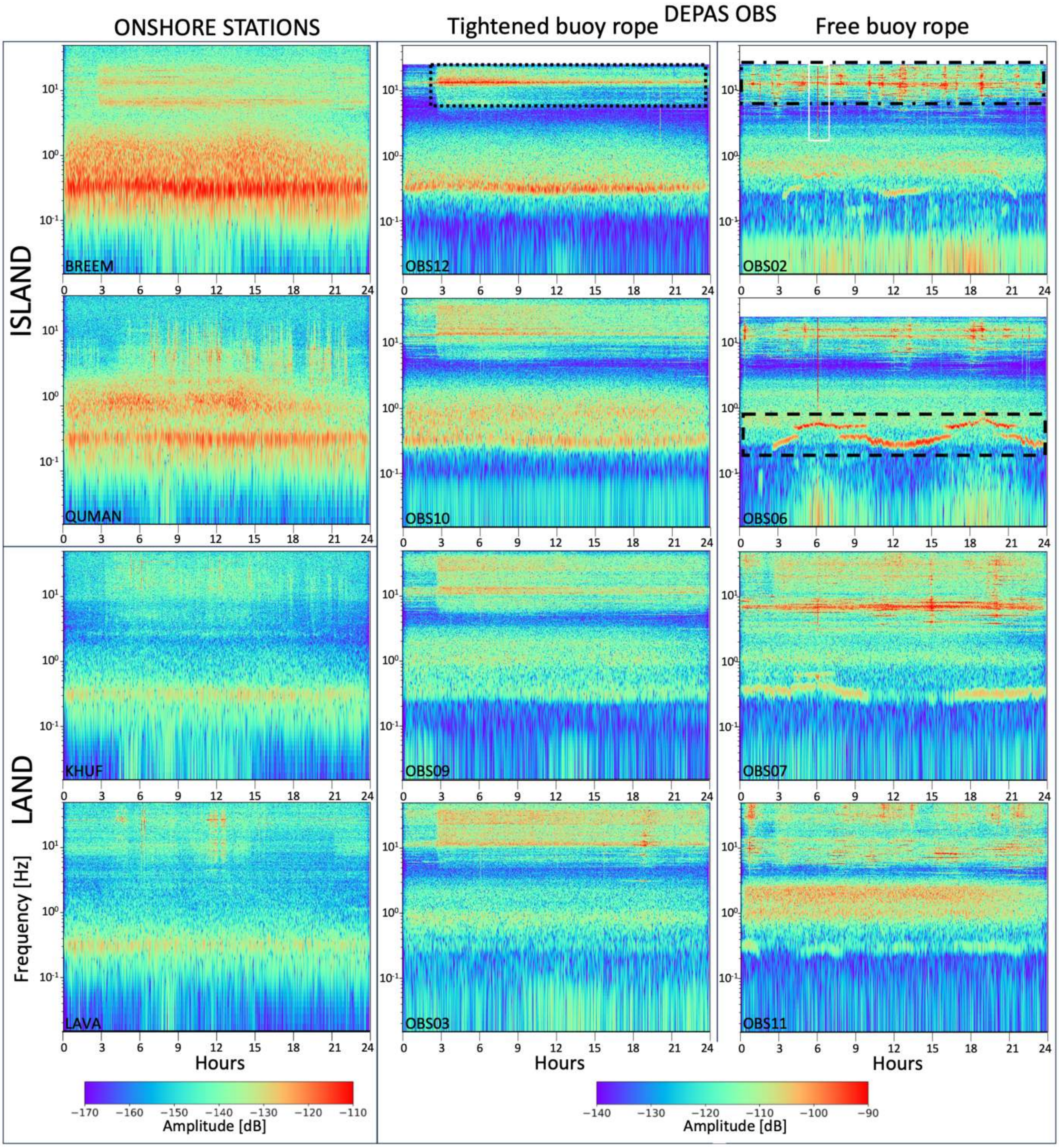


Figure 11 One-day spectrograms with examples of the notable signals detected by the ZAFRAN network. The left column shows the spectrograms for onshore stations (island stations in the first two rows and land stations in the last two rows). The central and right columns show the spectrograms for 8 of the DEPAS OBS. The central column contains spectrograms for OBS with tightened buoy rope. The right column contains spectrograms for OBS with a buoy rope free to strum in the water column. The dashed-line box on the spectrogram of OBS12 highlights an hours-long high-frequency signal (also evident in BREEM, KHUF, OBS10, OBS09, OBS03, OBS07). The dot-dashed-line box on the spectrogram of OBS02 highlights the high-frequency noise generated by the rope free to strum (also evident in all the other OBS with the free rope). The white box on the spectrogram of OBS02 highlights an example of a local earthquake (also visible in the spectrogram of OBS06). The dashed-line box on the spectrogram of OBS06 highlights a long-period tide-modulated signal (evident also in the other free buoy rope OBS). All spectrograms are calculated instrument-corrected acceleration waveforms of the horizontal component 1/N for March 1, 2022.

find that the free rope affects the dataset at frequencies higher than 5 Hz (0.2 s).

The second type of noise, visible only on the spectrograms of stations with the free rope, is a bi-daily tidal modulated signal with periods of 2-7 s. OBS06 exhibits the highest amplitudes of this signal (see the signal enclosed in the dashed-line box in Figure 11). Schlindwein et al. (2018) and Hannemann et al. (2016) found a similar signal in less than half of the DEPAS OBSs used in their deployments and as they could not identify the source, they assumed it to be of instrumental origin.

4.3 Hybrid local earthquake

We conduct a preliminary analysis of a local earthquake visually detected in the spectrograms of Figure 7 (see red box "LOC") that occurred on June 30, 2022. A 150s-long time window, containing this earthquake, is shown in Figure 12a. After manually picking P and S arrivals, we calculated the epicenter using HypoInverse code (Klein, 2002). The hypocentral depth is kept fixed at 5 km because it strongly depends on the velocity model, which still needs to be optimized for the ZFZ. Our preliminary calculations for earthquake location and magnitude suggest that this event occurred in the southern part of the network (see location indicated by the yellow circle in Figure 1) with a magnitude of M_L 3.4. This is likely the strongest local earthquake recorded by the ZAFRAN network since no stronger signals are seen in Figure 7 at high frequency. Not having access to the recent Saudi national catalog and being too small to be detected by the global networks, we cannot benchmark our location and magnitude estimate.

Qualitative waveform differences between stations may already reveal important information for future works. For example, waveforms (highpass filtered at 0.1 Hz) recorded by OBS10, KHUF, and OBS12 (Figure 12a) have smaller amplitudes after the P arrivals compared to waveforms in the other stations, probably due to weaker scattering effects. One hypothesis is that these differences are related to the site effects and that OBSs displaying waveforms with less scattering effects are located on thinner or no sedimentary cover (evaporites and/or loose sediments). Although the actual distribution and thickness of these sedimentary materials is not known in the ZFZ, this is consistent with the locations of OBS10, located in the deepest part of the ZFZ where the salt coverages seem to be limited, and OBS12, located on the flanks of Mabahiss Mons (Fittipaldi et al., 2024). While these site-effects may limit the ability to accurately pick the onset of the body waves both for location and focal mechanism calculations, they could be used for retrieving the shallow earth structure.

Waveforms for the same earthquake show a clear T-Phase at some of the stations (see boxes on waveforms of OBS04, OBS07, OBS09 OBS11, BREEM, and OBS12 in Figure 12a). T-phases are generated by the seismic-acoustic conversion and travel along the minimum sound velocity layer in the ocean (e.g., Okal, 2008). In our waveforms, the T-phase starts to be visible at stations farther than 70 km from the source (for closer stations, it is probably contaminated by the S coda). However, some stations at large distances do not show the T-phase. The ZAFRAN dataset has thus the potential to provide information on T-phase generation and its relation with the sea bottom topography and seawater layering in the Red Sea.

The spectral content of this earthquake is particularly intriguing because of its low-frequency amplitudes. Figures 12b-e show the spectra and the spectrograms for the three closest stations (OBS01, OBS02, and SOUTH, located between 17 and 28 km from the epicenter). The low-frequency content is also highlighted by waveforms for the three closest stations, low-pass filtered at 1 Hz (Figure 12f). This local earthquake is a combination of short-duration high-frequency and long-duration low-frequency events with the low-frequency being recorded first (Figure 12c-e and gray boxes in Figure 12c). These features classify it as a hybrid earthquake and are usually associated with fluid movements and/or conduit resonance due to volcanic or hydrothermal activities (e.g., Chouet, 1996; Neuberg et al., 2000; Leva et al., 2022). However, hybrid events are usually characterized by high-frequency onsets that generate low-frequency resonance of the rocks hosting the fluids (Neuberg et al., 2006). Other studies pointed out that the low-frequency content may be due to deep source or complex path effects (Harrington and Brodsky, 2007; Leva et al., 2022). In the case of the earthquake recorded by the ZAFRAN network, the onset of the low-frequencies is earlier than the onset of the high frequencies, not fitting well with the model that includes the resonance effects after the high-frequency rupture.

Comparison of spectra of signals recorded by the DEPAS OBSs (OBS01 and OBS02) and the Fugro OBS (SOUTH) reveals that the DEPAS OBSs experience resonance at 6 and 10.5 Hz following for at least 150 s after the arrival of the body waves (see Figure 12c-e). SOUTH, instead, does not show these signals. A similar signal at 6 Hz was observed in the DEPAS OBSs by Essing et al. (2021b) and attributed to the vibrations of the radio antenna.

4.4 Teleseismic waveforms

We conduct a qualitative assessment of waveform quality in the long-period range using a 2.5-hour-long time window that displays two Mw 6.7 teleseismic earthquakes. These earthquakes are detectable in spectrograms shown in Figure 7 (see the red box "TL") and are listed in all global earthquake catalogs. They occurred approximately at the same epicentral distance of 73^o but had different back-azimuths (Figure 13a). Their origin times differ by an hour. Band-passed waveforms between 5 and 100 s are displayed in Figure 13b.

Despite the high self-noise of the CMG-40T-OBS sensors in the DEPAS OBSs, the waveform quality for these two earthquakes is generally good, particularly on the vertical components. However, the noise in the horizontal components of some DEPAS OBSs makes it difficult to recognize the arrivals of different phases (Figure 13b, gray waveforms). This is the case for OBS01, OBS02, OBS06, and OBS11. Although the overall median noise

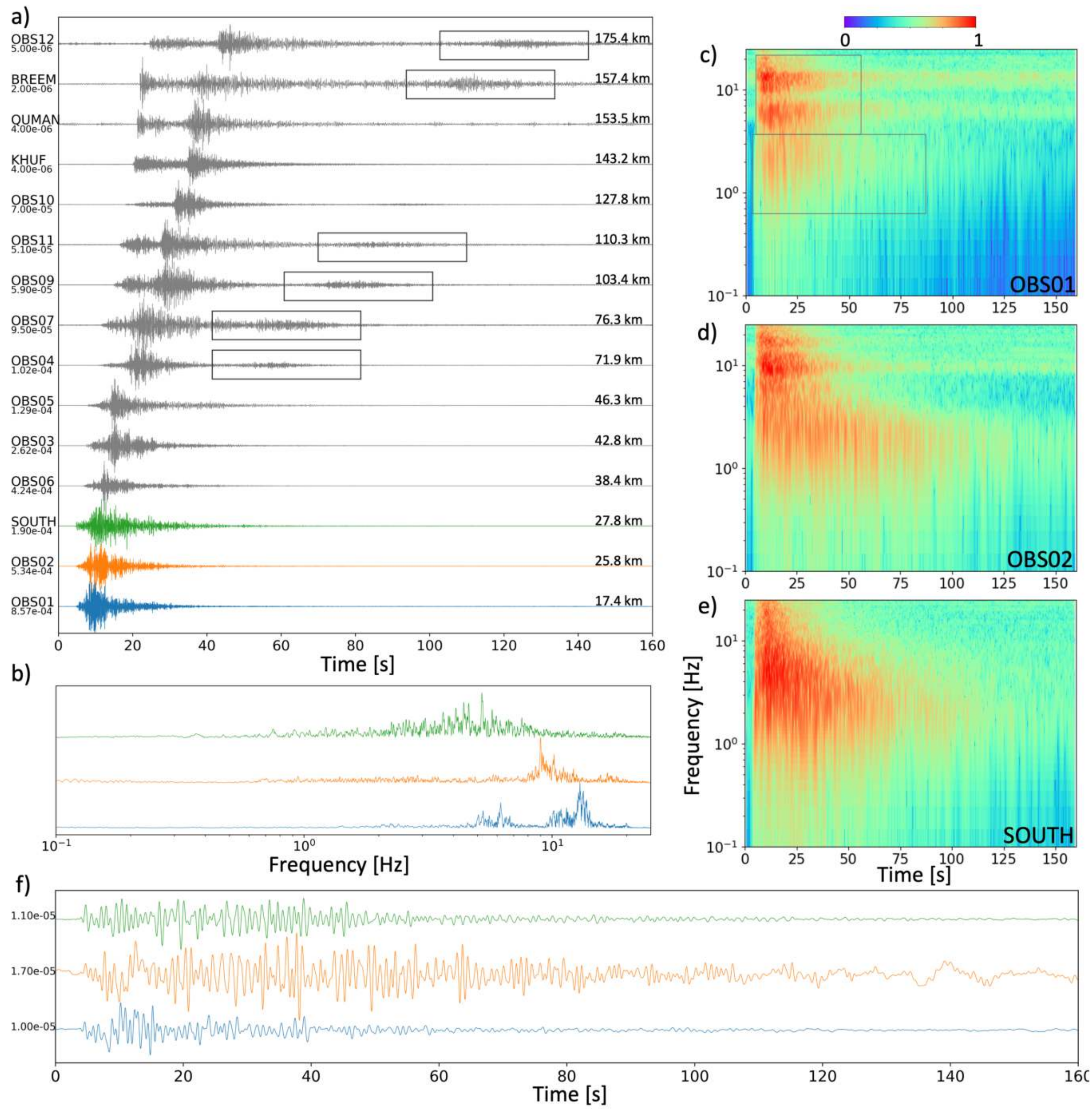


Figure 12 The M$_L$ 3.4 local hybrid earthquake. a) Vertical-component velocity seismograms and highpass filtered at 0.1 Hz. The maximum amplitude in m/s for each waveform is shown below the station's name. Possible T-phases are enclosed in the gray boxes. b) Normalized power spectra of the waveforms in stations OBS01, OBS02, and SOUTH (time window and colors as in a)). c) Normalized spectrograms for the station OBS01. Gray boxes highlight the short-duration high-frequency event and the long-duration low-frequency event. d) and e) as in c) but for OBS02 and SOUTH, respectively. f) Vertical-component velocity seismograms and low-pass filtered at 1 Hz. Colors and units as in a). The location of the earthquake is shown in Figure 1.

level of component 2 of OBS06 is not significantly higher than that of most other DEPAS OBSs (Figure 6), Figure 9 shows that in March 2022, the noise level on component 2 of OBS06 was almost 10 dB higher than the full deployment median. Instead, the overall median noise levels for horizontal component 2 of OBS01, OBS02, and OBS11 are generally higher than in the other stations, so their low waveform quality for these events is expected. The horizontal component waveforms of OBS03 began to exhibit noise at the end of the second earthquake, suggesting a possible temporary malfunction.

The land stations recorded the events excellently across all components (Figure 13b). The Fugro OBS SOUTH recorded the two events in the vertical component with a quality comparable to that of the onshore stations and in the horizontal component with a quality comparable to the average of the DEPAS OBSs.

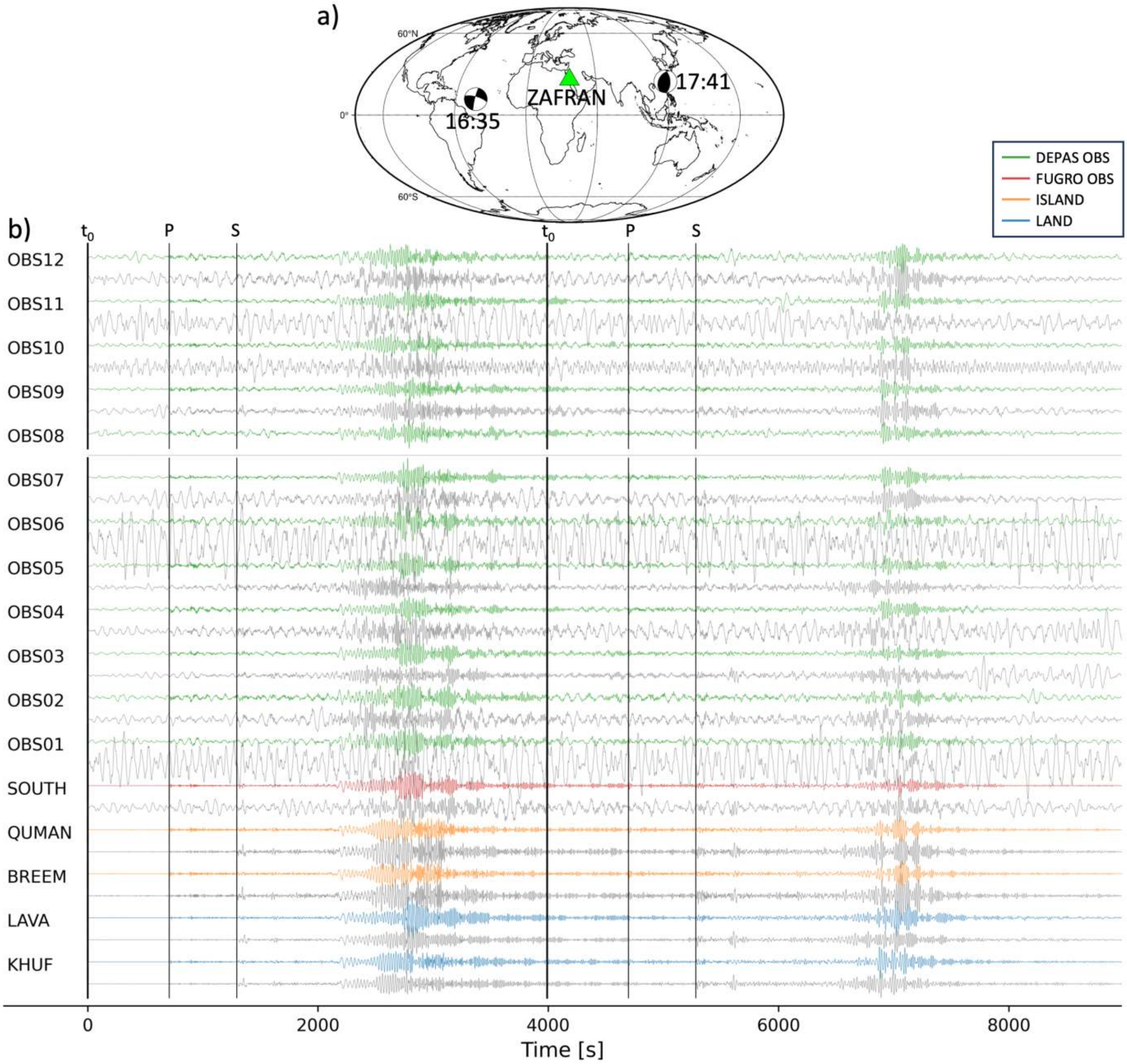


Figure 13 Example of two teleseismic earthquakes. a) Location and focal mechanisms for two M_w 6.7 teleseismic earthquakes that occurred on 22 March 2022 (t_0 displayed next to the focal mechanism). These events are recorded at approximately the same epicentral distance of 73^o. b) 5-100 s bandpass filtered velocity waveforms for the teleseismic earthquakes in a). Vertical seismograms are color-coded as in the legend with the name of the station displayed on the left. 2/E component seismograms are displayed in gray after the corresponding vertical component seismogram. Thick vertical black lines mark the earthquake origin times and thin black lines mark the P and S wave arrivals based on the IASP91 Earth model (Kennet, 1991). Event information is extracted from the Global CMT Catalog (https://www.globalcmt.org, last accessed 5 April 2023).

5 Discussion and conclusions

In this study, we introduce the first OBS deployment in the Red Sea, targeting the seismic activity and the lithospheric structure of the ZFZ. To establish a foundation for seismological investigations incorporating data from the ZAFRAN network, we perform comprehensive data quality control and emphasize notable signals recorded in the ZFZ.

For local seismicity studies, it is important to consider site effects arising from sedimentary and salt coverage since these factors may impact automated methods for determining arrival times and onset polarities on waveforms of local earthquakes. Moreover, in the short-period range, the dataset experiences higher noise levels compared to most global locations (e.g., Stähler et al., 2016, 2018; Hilmo and Wilcock, 2020; Carchedi et al., 2022; Kim et al., 2023; Zhang et al., 2023); this may be attributed to the regular passage of ships, intense seismic or tremor-like activities related to volcanic or hydrothermal phenomena, as suggested by the detection of the hybrid earthquake, or the slow movements of salt coverages resembling glacial creep (Podolskiy et al., 2021). Further disturbances in the short period are due to the strumming of the rope (frequencies larger than 10 Hz) and vibration of other OBS components (frequencies between 6 and 10 Hz).

Several recent studies attempted to model the harmonic noise generated by the head-buoy and rope strumming in the water columns of the DEPAS OBSs (Stähler et al., 2018; Essing et al., 2021b; Corela et al., 2022); however, this was observed at frequencies between 1 and 10 Hz. For example, Stähler et al. (2018) modeled the strumming of the rope and head buoy as it is usually implemented in the Lobsters of the DEPAS pool, including in our deployment. They found that the fundamental frequency of the DEPAS setup's rope is around 1 Hz, roughly corresponding to the shedding frequency of vortices generated by currents at approximately 0.1 m/s (calculated using the formula $f_{vort} = 10.5 \times v$ Hz, with v representing current velocity). Overtones were observed up to 10 Hz. Even considering the potentially higher salinity and temperature of Red Sea water compared to the values used by Stähler et al. (2018), the derivation of the f_{vort} does not change significantly. Consequently, the higher frequencies observed in our dataset suggest sea bottom currents of about 100 cm/s. Currents exceeding 100 cm/s have been documented in selected locations worldwide, such as the Gulf of Mexico and the Strait of Gibraltar, as reported by Shanmugam (2021). Unfortunately, specific sea bottom current data for the Red Sea are unavailable. However, ROV images of the deep Red Sea show that sea-bottom currents are usually weak and not able to blow away light bacterial mats (van der Zwan et al., 2023). These observations imply that either the signal we observe is unrelated to the strumming rope (in this case their occurrence in the OBS with the free rope is by chance only and due to specific locations instead) or the phenomenon interacting with the rope is not a sea bottom current.

In the medium-period range, relevant for example for ambient noise tomography, our analysis of the correlation with the oceanic and atmospheric parameters indicates that local noise sources from wind and waves may adversely affect the quality of noise cross-correlations. Consequently, we recommend utilizing cross-correlations calculated during days of calm local sea states, emphasizing the importance of selecting or weighing cross-correlations using datasets like ERA. The analyses of correlations presented here are extremely important in defining the period range (0.2 - 10 s) of the influence of wind and waves on ambient noise.

In addition, our analyses contribute to a better understanding of a second peak within the medium-period range, occurring at shorter periods than the secondary microseismic peak. The presence of two distinct peaks within the secondary microseismic band is commonly interpreted as occurred due to sources at different distances with the peak at a shorter period generated by local sea conditions and the peak at a longer period due to farther oceanic sources (e.g., Bromirski et al., 2005; Zhang et al., 2023). In our study, the presence of the shorter period peak in the noise of the OBSs and island stations and the lack of the peak in the land stations (see comparison in Figure 5) supports the fact that this shorter period peak is due to local sources whose energy dissipates rapidly inland.

On the other hand, the second peak at a shorter period was also observed in stations located far from the coastline and attributed to and used to retrieve subsurface structure (Parisi et al., 2020; Guo et al., 2021). For example, Kim et al. (2023) also found that the thickness of sediments below the OBSs attenuates (amplifies) periods shorter (longer) than 2 s. In this study, we find that the noise level in the medium-period range correlates well with the sea state, and the correlation functions ρ_{10v} and ρ_{swh} have two peaks at frequencies similar to ones of the noise levels in the medium-period range. Therefore, we would expect that all stations with ρ_{10v} and ρ_{swh} with double peaks should have a double peak noise level, that is, the stations should belong to class A or B. However, we observe that even stations of other classes (no second peak or a large peak including the frequency from the first and second peak), such as OBS12 belonging to class D, have a ρ_{10v} and ρ_{swh} with double peak (see Figure 10). We believe that the shape of the noise level of these stations is due to the effects of the Earth's structure below the station that amplifies some frequencies between the two peaks. Also, the fact that the shape/category of the ZAFRAN stations does not vary in time (Figure 9) further supports the effect of local structure, in addition to the local sources.

The factor significantly affecting the performance of the DEPAS OBS at long-period is the high self-noise levels of the Güralp CMG-40T-OBS sensors, limiting observations in the long-period range (T>10s) in all components (Figure 6). In contrast, the Fugro OBSs, equipped with a Nanometrics Trillium compact sensor, performed as well as the land stations on the vertical component (Figure 6). Noise levels on the horizontal components were comparable in the DEPAS and Fugro OBSs. These findings are consistent with observations made by Stähler et al. (2018) when comparing the two types of sensors installed on Lobsters (the DEPAS OBS configuration). The differences observed in the performances of the OBSs at long periods directly reflect on the quality of waveforms for teleseismic earthquakes. The example in Figure 13 in fact shows the overall good quality of the network with some limitations in some horizontal components of the DEPAS OBS. However, the clear identifications of body and surface waves suggest that the ZAFRAN dataset can be included in global seismology studies for earthquakes of magnitude at least M_w6.7, or less if only the Fugro OBSs are used.

The findings presented in this study hold significance for forthcoming research relying on the ZAFRAN dataset. They bear importance for geological and oceanographic investigations in the Red Sea, as well as for the seismological communities engaged in refining their understanding and use of OBSs.

Acknowledgements

We would like to acknowledge the “German Instrument Pool for Amphibian Seismology (DEPAS)”, hosted by the Alfred Wegener Institute Bremerhaven, for providing the OBSs and Fugro for providing access to their OBS dataset. We thank the Coastal and Marine Resource

Core lab of KAUST and in particular Abdullah Mandili, David Atienza, and the crews of the R/V Thuwal and R/V Azizi. We are also grateful to the legal and logistic teams at KAUST and GEOMAR. Moreover, we thank Red Sea Global for supporting our project and providing field support. Lastly, we are grateful to the entire SEISMICA team for their efforts in facilitating the publication of this manuscript, with a special thank to the invaluable feedback provided by the Editor Stephen P. Hicks and two anonymous reviewers, whose comments improved the quality of our work. This research was funded by the KAUST Competitive Research Grant URF/1/4076-01-01.

Data and code availability

Waveforms and instrument responses will be publicly available from the GEOFON data center with network code 5Q after the embargo period (3 years from the end of the deployment). Figure 1 was made with QGIS; Figure 6 with ObsPy (Krischer et al., 2015) and matplotlib; Figures 7,11,12 and 13 with ObsPy. ERA5 datasets are available from ECMWF (European Centre for Medium-Range Weather Forecasts, Hersbach et al., 2023, last access 29 August 2023).

Competing interests

The authors have no competing interests.

References

Alfred-Wegener-Institut Helmholtz-Zentrum für Polar-und Meeresforschung et al., . DEPAS (Deutscher Geräte-Pool für amphibische Seismologie): German Instrument Pool for Amphibian Seismology. *Journal of large-scale research facilities*, 3(A122), 2017. doi: 10.17815/jlsrf-3-165.

Almalki, K. A., Betts, P. G., and Ailleres, L. The Red Sea – 50years of geological and geophysical research. *Earth-Science Reviews*, 147:109–140, 2015. doi: 10.1016/j.earscirev.2015.05.002.

Ardhuin, F. Large-Scale Forces Under Surface Gravity Waves at a Wavy Bottom: A Mechanism for the Generation of Primary Microseisms. *Geophysical Research Letters*, 45(16):8173–8181, 2018. doi: 10.1029/2018GL078855.

Ardhuin, F., Rawat, A., and Aucan, J. A numerical model for free infragravity waves: Definition and validation at regional and global scales. *Ocean Modelling*, 77:20–32, 2014. doi: 10.1016/j.ocemod.2014.02.006.

Augustin, N., van der Zwan, F. M., Devey, C. W., and Brandsdóttir, B. 13 million years of seafloor spreading throughout the Red Sea Basin. *Nature Communications*, 12(1), 2021. doi: 10.1038/s41467-021-22586-2.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., and Wassermann, J. ObsPy: A Python Toolbox for Seismology. *Seismological Research Letters*, 81(3):530–533, 05 2010. doi: 10.1785/gssrl.81.3.530.

Blanck, H., Jousset, P., Hersir, G. P., Ágústsson, K., and Ólafur G. Flóvenz. Analysis of 2014–2015 on- and off-shore passive seismic data on the Reykjanes Peninsula, SW Iceland. *Journal of Volcanology and Geothermal Research*, 391:106548, 2020. doi: 10.1016/j.jvolgeores.2019.02.001.

Braunmiller, J., Nabelek, J., and Ghods, A. Sensor Orientation of Iranian Broadband Seismic Stations from P-Wave Particle Motion. *Seismological Research Letters*, 91(3):1660–1671, 04 2020. doi: 10.1785/0220200019.

Bromirski, P. D., Duennebier, F. K., and Stephen, R. A. Mid-ocean microseisms. *Geochemistry, Geophysics, Geosystems*, 6(4), 2005. doi: 10.1029/2004GC000768.

Carchedi, C. J. W., Gaherty, J. B., Webb, S. C., and Shillington, D. J. Investigating Short-Period Lake-Generated Microseisms Using a Broadband Array of Onshore and Lake-Bottom Seismometers. *Seismological Research Letters*, 93(3):1585–1600, 02 2022. doi: 10.1785/0220210155.

Chouet, B. A. Long-period volcano seismicity: Its source and use in eruption forecasting. *Nature*, 380(6572):309 – 316, 1996. doi: 10.1038/380309a0.

Coleman, R. G. and McGuire, A. V. Magma systems related to the Red Sea opening. *Tectonophysics*, 150(1):77–100, 1988. doi: 10.1016/0040-1951(88)90296-X.

Corela, C., Loureiro, A., Duarte, J. L., Matias, L., Rebelo, T., and Bartolomeu, T. The OBS noise due to deep ocean currents. *Natural Hazards and Earth System Sciences Discussions*, 2022:1–21, 2022. doi: 10.5194/nhess-2022-196.

Crawford, W. C. and Webb, S. C. Identifying and Removing Tilt Noise from Low-Frequency (<0.1 Hz) Seafloor Vertical Seismic Data. *Bulletin of the Seismological Society of America*, 90(4): 952–963, 08 2000. doi: 10.1785/0119990121.

Delaunay, A., Baby, G., Fedorik, J., Afifi, A. M., Tapponnier, P., and Dyment, J. Structure and morphology of the Red Sea, from the mid-ocean ridge to the ocean-continent boundary. *Tectonophysics*, 849:229728, 2023. doi: 10.1016/j.tecto.2023.229728.

Doran, A. K. and Laske, G. Ocean-Bottom Seismometer Instrument Orientations via Automated Rayleigh-Wave Arrival-Angle Measurements. *Bulletin of the Seismological Society of America*, 107 (2):691–708, 01 2017. doi: 10.1785/0120160165.

El-Isa, Z. Seismicity and seismotectonics of the Red Sea Region. *Arabian Journal of Geosciences*, 8(10):8505 – 8525, 2015. doi: 10.1007/s12517-015-1819-2.

Essing, D., Schlindwein, V., Schmidt-Aursch, M. C., Hadziioannou, C., and Stähler, S. C. Characteristics of Current-Induced Harmonic Tremor Signals in Ocean-Bottom Seismometer Records. *Seismological Research Letters*, 92(5):3100–3112, 04 2021a. doi: 10.1785/0220200397.

Essing, D., Schlindwein, V., Schmidt-Aursch, M. C., Hadziioannou, C., and Stähler, S. C. Characteristics of Current-Induced Harmonic Tremor Signals in Ocean-Bottom Seismometer Records. *Seismological Research Letters*, 92(5):3100–3112, 04 2021b. doi: 10.1785/0220200397.

Fittipaldi, M., Trippanera, D., Augustin, N., van der Zwan, Parisi, L., and Jónsson, S. Mabahiss Deep in the Northern Red Sea: New geomorphological insights from high-resolution bathymetric mapping . *Submitted to Geomorphology*, 2024.

Geissler, W. H., Matias, L., Stich, D., Carrilho, F., Jokat, W., Monna, S., IbenBrahim, A., Mancilla, F., Gutscher, M.-A., Sallarès, V., and Zitellini, N. Focal mechanisms for sub-crustal earthquakes in the Gulf of Cadiz from a dense OBS deployment. *Geophysical Research Letters*, 37(18), 2010. doi: 10.1029/2010GL044289.

Gualtieri, L., Stutzmann, E., Capdeville, Y., Ardhuin, F., Schimmel, M., Mangeney, A., and Morelli, A. Modelling secondary microseismic noise by normal mode summation. *Geophysical Journal International*, 193(3):1732–1745, 04 2013. doi: 10.1093/gji/ggt090.

Guo, Z., Huang, Y., and Aydin, A. Double-Frequency Microseisms on the Thick Unconsolidated Sediments in Eastern and Southeastern Coasts of United States: Sources and Applications on Seismic Site Effect Evaluation. *Journal of Earth Science*, 32(5):1190

– 1201, 2021. doi: 10.1007/s12583-021-1426-y.

Hamieh, A., Rowaihy, F., Al-Juaied, M., Abo-Khatwa, A. N., Afifi, A. M., and Hoteit, H. Quantification and analysis of CO2 footprint from industrial facilities in Saudi Arabia. *Energy Conversion and Management: X*, 16:100299, 2022. doi: 10.1016/j.ecmx.2022.100299.

Hannemann, K., Krüger, F., Dahm, T., and Lange, D. Oceanic lithospheric S-wave velocities from the analysis of P-wave polarization at the ocean floor. *Geophysical Journal International*, 207 (3):1796–1817, 09 2016. doi: 10.1093/gji/ggw342.

Harrington, R. M. and Brodsky, E. E. Volcanic hybrid earthquakes that are brittle-failure events. *Geophysical Research Letters*, 34 (6), 2007. doi: 10.1029/2006GL028714.

Heleno, S. I., Faria, B. V., Bandomo, Z., and Fonseca, J. F. Observations of high-frequency harmonic tremor in Fogo, Cape Verde Islands. *Journal of Volcanology and Geothermal Research*, 158 (3):361–379, 2006. doi: 10.1016/j.jvolgeores.2006.06.018.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N. ERA5 hourly data on single levels from 1940 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*, 2023. doi: 10.24381/cds.adbb2d47.

Hilmo, R. and Wilcock, W. S. D. Physical Sources of High-Frequency Seismic Noise on Cascadia Initiative Ocean Bottom Seismometers. *Geochemistry, Geophysics, Geosystems*, 21(10): e2020GC009085, 2020. doi: 10.1029/2020GC009085.

Janiszewski, H. A., Gaherty, J. B., Abers, G. A., Gao, H., and Eilon, Z. C. Amphibious surface-wave phase-velocity measurements of the Cascadia subduction zone. *Geophysical Journal International*, 217(3):1929–1948, 01 2019. doi: 10.1093/gji/ggz051.

Janiszewski, H. A., Eilon, Z., Russell, J. B., Brunsvik, B., Gaherty, J. B., Mosher, S. G., Hawley, W. B., and Coats, S. Broad-band ocean bottom seismometer noise properties. *Geophysical Journal International*, 233(1):297–315, 11 2022. doi: 10.1093/gji/ggac450.

Kennet, B. L. N. IASPEI 1991 SEISMOLOGICAL TABLES. *Terra Nova*, 3(2):122–122, 1991. doi: 10.1111/j.1365-3121.1991.tb00863.x.

Kim, T., Park, J., Ko, J., Oh, S., Witek, M., Chang, S., Lee, S., Kim, Y., Utada, H., Kawakatsu, H., Shiobara, H., Isse, T., Takeuchi, N., and Sugioka, H. Characteristics of Background Noise in the Oldest-1 Array Deployed on the Oldest Part of the Pacific Plate. *Bulletin of the Seismological Society of America*, 113(4):1772–1793, 04 2023. doi: 10.1785/0120220215.

Klein, F. W. User's guide to HYPOINVERSE-2000, a Fortran program to solve for earthquake locations and magnitudes. *USGS, Open-File Report 2002-171*, 2002.

Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C., and Wassermann, J. ObsPy: a bridge for seismology into the scientific Python ecosystem. *Computational Science & Discovery*, 8(1):014003, may 2015. doi: 10.1088/1749-4699/8/1/014003.

Leva, C., Rümpker, G., and Wölbern, I. Multi-array analysis of volcano-seismic signals at Fogo and Brava, Cape Verde. *Solid Earth*, 13(8):1243–1258, 2022. doi: 10.5194/se-13-1243-2022.

Marshak, S., Bonatti, E., Brueckner, H., and Paulsen, T. Fracture-zone tectonics at Zabargad Island, Red Sea (Egypt). *Tectonophysics*, 216(3):379–385, 1992. doi: 10.1016/0040-1951(92)90407-W.

McNamara, D. E. and Buland, R. P. Ambient Noise Levels in the Continental United States. *Bulletin of the Seismological Society of America*, 94(4):1517–1527, 08 2004. doi: 10.1785/012003001.

Naranjo, D., Parisi, L., Jónsson, S., Jousset, P., Werthmüller, D., and Weemstra, C. Ocean Bottom Seismometer Clock Correction using Ambient Seismic Noise. *Seismica*, 3(1), Jan. 2024. doi: 10.26443/seismica.v3i1.367.

Neuberg, J., Luckett, R., Baptie, B., and Olsen, K. Models of tremor and low-frequency earthquake swarms on Montserrat. *Journal of Volcanology and Geothermal Research*, 101(1):83–104, 2000. doi: 10.1016/S0377-0273(00)00169-4.

Neuberg, J., Tuffen, H., Collier, L., Green, D., Powell, T., and Dingwell, D. The trigger mechanism of low-frequency earthquakes on Montserrat. *Journal of Volcanology and Geothermal Research*, 153(1):37–50, 2006. doi: 10.1016/j.jvolgeores.2005.08.008.

Okal, E. A. The generation of T waves by earthquakes. volume 49 of *Advances in Geophysics*, pages 1–65. Elsevier, 2008. doi: 10.1016/S0065-2687(07)49001-X.

Parisi, L., Stanistreet, I., Njau, J., Schick, K., Toth, N., and Mai, P. M. Seismological Investigations in the Olduvai Basin and Ngorongoro Volcanic Highlands (Western Flank of the North Tanzanian Divergence). *Seismological Research Letters*, 91(6):3286–3303, 09 2020. doi: 10.1785/0220200111.

Peterson, J. Observations and Modeling of Seismic Background Noise. *U.S. Geological Survey open-file report*, Albuquerque, N.M:93–3222, 1993. doi: 10.1016/j.ocemod.2014.02.006.

Podolskiy, E. A., Murai, Y., Kanna, N., and Sugiyama, S. Ocean-bottom and surface seismometers reveal continuous glacial tremor and slip. *Nature Communications*, 12(1), 2021. doi: 10.1038/s41467-021-24142-4.

Rehman, F., Alamri, A. M., El-Hady, S. M., Harbi, H. M., and Atef, A. H. Seismic hazard assessment and rheological implications: a case study selected for cities of Saudi Arabia along the eastern coast of Red Sea. *Arabian Journal of Geosciences*, 10(24), 2017. doi: 10.1007/s12517-017-3325-1.

Schlindwein, V., Krüger, F., and Schmidt-Aursch, M. Project KNIPAS: DEPAS ocean-bottom seismometer operations in the Greenland Sea in 2016-2017, 2018. doi: 10.1594/PANGAEA.896635.

Shanmugam, G. Chapter 8 - Bottom currents. In Shanmugam, G., editor, *Mass Transport, Gravity Flows, and Bottom Currents*, pages 309–375. Elsevier, 2021. doi: 10.1016/B978-0-12-822576-9.00008-4.

Stutzmann, E., Roult, G., and Astiz, L. GEOSCOPE Station Noise Levels. *Bulletin of the Seismological Society of America*, 90(3): 690–701, 06 2000. doi: 10.1785/0119990025.

Stähler, S., Sigloch, K., Hosseini, K., Crawford, W., Barruol, G., Schmidt-Aursch, M., Tsekhmistrenko, M., Scholz, J.-R., Mazzullo, A., and Deen, M. Performance report of the RHUM-RUM ocean bottom seismometer network around la Réunion, western Indian Ocean. *Advances in Geosciences*, 41:43 – 63, 2016. doi: 10.5194/adgeo-41-43-2016.

Stähler, S. C., Schmidt-Aursch, M. C., Hein, G., and Mars, R. A Self-Noise Model for the German DEPAS OBS Pool. *Seismological Research Letters*, 89(5):1838–1845, 08 2018. doi: 10.1785/0220180056.

Trabattoni, A., Barruol, G., Dréo, R., and Boudraa, A. Ship detection and tracking from single ocean-bottom seismic and hydroacoustic stations. *The Journal of the Acoustical Society of America*, 153(1):260–273, 01 2023. doi: 10.1121/10.0016810.

van der Zwan, F. M., Augustin, N., Petersen, S., Altalhi, S. M., Schultz, J., Peixoto, R. S., Follmann, J., Anker, A., Benzoni, F., Garcia Paredes, E. R., Al Malallah, M., Shepard, L., Ouhssain, M., Jägerup, S. B., Jones, B. H., and Rosado, A. S. Widespread diffuse venting and large microbial iron-mounds in the Red Sea. *Communications Earth and Environment*, 4(1), 2023. doi: 10.1038/s43247-023-01169-7.

Wilcock, W. S. D. Tracking fin whales in the northeast Pacific Ocean with a seafloor seismic network. *The Journal of the Acoustical Society of America*, 132(4):2408–2419, 10 2012. doi: 10.1121/1.4747017.

Zhang, H., Schmidt-Aursch, M. C., Geissler, W. H., and Xing, J. Characteristics of the Oceanic Ambient Seismic Noise Around Tristan da Cunha in the South Atlantic From OBS Data. *Journal of Geophysical Research: Solid Earth*, 128(6):e2022JB025884, 2023. doi: 10.1029/2022JB025884.