

Dear Editor (Christos Evangelidis) and Reviewers,

We sincerely thank you for your time, careful reading, and constructive feedback. We have revised the manuscript thoroughly in response to all comments. The revised version has also undergone comprehensive language editing to improve clarity and readability.

Below, we provide a detailed, point-by-point response to each reviewer comment. All changes have been incorporated into the manuscript, and a marked-up version highlighting modifications has been provided.

The manuscript has been improved in the following areas:

- **Language and clarity:** The entire manuscript has been carefully proofread and revised to correct grammatical issues and improve readability.
- **Motivation and scope:** The Introduction has been significantly expanded to clarify the scientific and operational motivation, particularly for applications in induced seismicity and aseismic regions.
- **Methodological clarification:**
 - Justification of template grid design, including depth constraints.
 - Explanation of the use of a combined P+S window versus separate phase windows.
 - Clarification of time-smearing and uncertainty handling across stations.
- **Interpretation improvements:**
 - Expanded discussion of trade-offs between detection sensitivity, location accuracy, and focal mechanism resolution.
 - Added explanation of limitations related to velocity model and cycle skipping.
- **Figures and references:**
 - Corrected and clarified figure captions and references.
 - Improved explanations of figure interpretation where needed.
- **Validation and comparison:**
 - Clarified catalog matching procedure and validation of detections.
 - Improved discussion of completeness and missed detections.
- **Manuscript structure:**
 - Improved section cross-referencing and logical flow.
 - Removal of incomplete code listing; clarified computational workflow in text.
 - Removal of detailed velocity model, replaced with references where appropriate

We believe the revisions have significantly improved the clarity, rigor, and impact of the manuscript. We appreciate the reviewers' and editor's efforts, which have helped strengthen this work.

We hope the revised manuscript meets the journal's standards and are happy to address any further comments.

Kind regards,
Eva Káldy, Jana Doubravová, Tomáš Fischer

Dear Eva Káldy, Jana Doubravová, Tomáš Fischer:

The review letters have also been included now.

My sincere apologies.

When you are ready to resubmit the revised version of your manuscript, please upload:

- A 'cleaned' version of the revised manuscript, without any markup/changes highlighted.
- A pdf version of the revised manuscript clearly highlighting changes/markup/edits.
- A 'response-to-reviewers' letter that shows your response to each of the reviewers' points, together with a summary of the resulting changes made to the manuscript.

Once I have read your revised manuscript and rebuttal, I will then decide whether the manuscript either needs to be sent to reviewers again, requires further minor changes, or can be accepted.

If you deem it appropriate, please check that the revised version of your manuscript recognises the work of the reviewers in the Acknowledgements section.

Please note that Seismica does not have any strict deadlines for submitting revisions, but naturally, it is likely to be in your best interest to submit these fairly promptly, and please let me know of any expected delays.

I wish you the best with working on the revisions. Please don't hesitate to contact me with any questions or comments about your submission, or if you have any feedback about your experience with Seismica.

Kind regards,

Christos Evangelidis
National Observatory of Athens
christos.evangelidis@seismica.org

Reviewer A:

This manuscript provides an interesting and potentially useful extension to matched-filter earthquake detection by extending the use of synthetic templates for matched-filter analysis to allow for uncertainties in synthetic arrival times and better characterisation of focal mechanisms. The authors demonstrate their new approach on a single day of data from the West Bohemia region and show that their method, when parameterised appropriately, had a very low false detection rate, although it does not meet the positive detection rates of other currently employed methods. Despite the performance of their new method not meeting the level of the methods they compare to, this manuscript remains a worthwhile contribution to the literature and should be considered suitable for publication with some amendments.

The main thing that needs to be corrected for this paper to be accepted is the quality of the writing. There are multiple grammatical errors throughout the manuscript that make many sentences difficult to understand, and several of my judgements on the validity of this research were hindered by these errors. Despite this, the science seems sound. Nevertheless, I have several suggestions, comments and questions that I think the authors should address.

1. Lines 195-201: The restriction of the template region in depth is likely necessary for efficiency, but I expect that in the unknown settings that the authors target this method at, the depth of

seismicity is unlikely to be known. I wonder how the authors would recommend setting up this grid if the depth of seismicity is unknown and how this would impact efficiency and location accuracy of detections.

- The Synthetic Template matching was primarily aimed at the induced seismicity detection. For that case the approximate area and depth of earthquakes is known.
- Incorporated in the text: In case of induced seismicity monitoring, the span of templates should cover the whole underground mass potentially influenced by human activity.

2. Line 224: It is common in matched-filters to define separate P and S windows to both reduce the noise in the template window and reduce the bias in the correlations due to the usually higher amplitude S arrival. It looks like you have chosen to use one window spanning the P and S phases (Fig. 2), although the text was not clear about this. Did you consider using separate P and S windows? I expect that this would help cope with differences in errors in P and S velocity models, rather than retaining a fixed S-P time.

- Yes, using separate P and S detection windows has benefits, but might be tricky at local settings as we plan for the induced seismicity. There the waveforms are quite basic in shape and if P and S were to be searched independently, it would cause a great deal of false triggers. Where I see the benefit of matched-filters separately on P and S is the relative location or re-location of already detected events.

- Modified: define *a single* time window relative to P and/or S wave.

When working in an area without recorded seismicity, the time difference between the real and synthetic wave arrival is not known, but it can be estimated from the uncertainty in the 1D velocity model. We estimate the expected error in the relative arrival time at the stations t_{err} to ~ 0.43 s. *In case of more complex shapes of synthetic waveforms the potential time difference between the real and synthetic wave arrival could be solved by having the match-filter defined separately in P and S window. For more local setting of induced seismicity, which we are aiming at, both the real and synthetic waveforms are basic in shape. There the individual P and S wave match-filter window would create a great deal of false triggers, which we decide to avoid by a single P+S match-filter window.*

3. Line 238: You state that the error is 0.43s to “some of the stations” then use this for all stations – please justify.

- In reality we would not know which of the stations is going to be affected, that is why we use the smearing for all the stations.
- Reformulated: We estimate the expected error in the relative arrival time at the stations t_{err} to ~ 0.43 s.

4. Given that you rely on five channels for detection this generally means that you only need similarity on two stations – did you investigate the number of stations for which the threshold was achieved for your detections? Could this be used as an additional quality check?

- You are likely referring to figure 4, which is used just for illustration. In the real setting the required number of components above the threshold is 9, which corresponds to min 3 stations that are considered minimum for location-triangulation. But yes, more settings were tested and this is rather conservative.

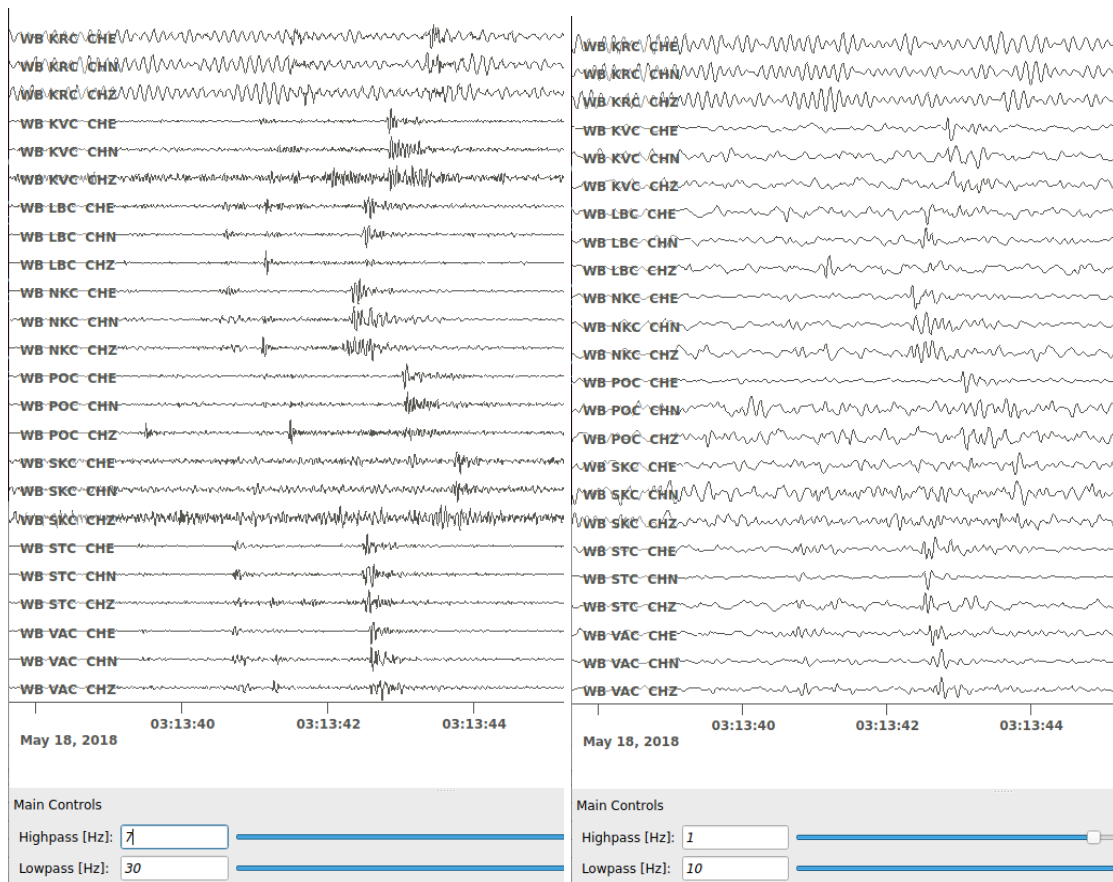
- Modified the text in figure 4: The detection threshold is set to 0.4, the required number of components above the threshold is set to 5 (only for this example on LBC, NKC and KVC station, *in real run it is set to 9*)...
 - Change of in Listing 1 (code): the `nr_components_above_threshold=9`
5. Given the reduced sensitivity to moveout in the templates, I was concerned that this method would provide lots of false detections for regional events. Have you tested this? Where there days when you know there was lots of regional seismicity that you could test your method against and see if it is able to discriminate between local and regional seismicity? It may be that the waveform shape is enough to discriminate between local and regional, but that has not been made clear here.
- During the test day there was also regional activity visible on WEBNET. None of the 3 visible and reported (Institute of Geophysics, 2024) was detected by synthetic template matching (xc0.4 nor xc0.35). On the other hand the detection with xc0.35 detected 11 distant or low-frequency events - none detected by PEPin or reported in the considered regional catalog...
 - Added: During this day there were also regional earthquakes (Institute of Geophysics, 2024), 3 of which from Lubin (Poland) are strong enough to be recorded in West Bohemia.
6. Fig 5 caption: reference for maximum curvature method is incorrect, these were not the authors who came up with the method.
- Thank you. Corrected for (Wiemer and Wyss, 2000)
7. Fig 5: Apologies, the scale is somewhat difficult to read without a ruler to hand: the orange dot (Synthetic template xc 0.35) a ML just below 0.0 has more events than the PEPiN and ANN catalogs, yet the completeness of these catalogs is reported as lower than the Synthetic. Are these false detections, or incorrect magnitudes?
- For the automatic magnitudes, I would not argue about its correctness and for small events the noise is affecting them significantly. But there are false unique events: in the bin M-0.2 – M0.0 there are 28 unique events, only 11 are real and local. I hope this is explained in section 4.2.2 where the correctness of unique events is discussed for match filter xc0.35 (typos corrected in that chapter).
8. Section 4.1: I am curious why the closest template to the location from other algorithms was not the final detection: why was the correlation of this template lower? What about that template makes it a less good fit?
- The templates closest to the other-algorithm locations also have high similarity, main uncertainty being in the depth. I did not look at the synthetic vs real waveforms through the various synthetic locations, but the difference in location is likely coming from the difference between the 1D-velocity model and real velocity field.
 - Look at the figure 2, where a real seismic event is overlaid by the synthetic one of similar focal mechanism and location: There is a time-lag at 4 of 8 stations visible at the North component (second column): the synthetics have S-wave arriving earlier at 3 stations (LBC, POC, KRC) North and East of the event location. This time lag implies that the interpreted and real location is likely to vary in N-E direction.

9. Section 4.1, Line 351: I suspect that the limited sensitivity to focal mechanism is a result of increased cycle-skipping. I note that in Fig, 4, similarity trace KVC.N the absolute maximum of the similarity is a negative number, implying that this trace would be better fit by the amplitude reversed version of this template. Because you are only considering positive peaks you are likely to make detections ~0.5 cycles out when the polarity of the waveform is reversed compared to the template. This would result in a trade-off between location (and origin-time) accuracy and focal mechanism sensitivity. I think this trade-off should be discussed.
- Yes, there is a tradeoff between location and focal mechanism - same as in the real FM inversion from amplitudes. The maximum value of similarity being negative in Fig. 4 (KVC.N, driven by the S-wave) is ok and if this was the only template detecting the real event, yes the interpreted FM would not be close to the correct one and the location would be also shifted (in this particular case 5 km apart as in the Fig 4 caption). But likely there is a template with S-wave amplitudes better matching the real FM and the location. The S-wave of templates vary significantly (Fig. 3). (In the Fig 4 example, the template detecting the real event with greatest similarity was only 600 m apart from the manual location.)
 - For the record: Half cycle skipping corresponds to error max 0.44 km at a station (synthetic half cycle ~1/16 s, velocity ~7 km/s). In our grid of synthetic templates (1 km step) such “picking” error due to phase skipping might be playing a role, if interpreted individually. But the smearing of similarities covers the skipping for a station of two, and assigns it to the same template as without phase skipping.
 - Overall the limited sensitivity to FM is more complex and is more related to the smearing the maximum similarity - it can trigger at each component at different cycles and still be summed up. That is positive for the detectability, less for the FM interpretation.
10. Section 4.2 – please use similar titles to the questions posed above to help the reader follow the text. You may also consider adding references to the relevant sections from the questions posed in lines 334-339.
- I add references to the questions. To make the questions titles , some are too long.
11. Line 361: I was surprised that you were able to simply match origin times between catalogues given the difference in origin times shown in section 4.1 for one event. How different did you allow origin times to be? Did you compare other characteristics to reliably associate matched-filter detections to other catalogue detections? It is not clear whether all detections made by the matched-filter were manually verified, or just the “new” detections. If only the new detections were verified, and only origin times were compared between catalogs, I suggest that more should be done to convince the reader that these detections are shared between catalogs.
- The match time was from -0.1 to + 4 seconds difference to template-matching catalog. This window was tested on a subset matched manually. When more other-catalog events fall into this window, the event closest to mean value (2.05 s) was chosen as the matched one. Only the origin times were considered.
 - Only all the unique detections were completely checked. I made it more clear in the text. “all additional events were visually examined and recognized real without a doubt.” All unique events $M > 0$ by PePin were also visually examined.

12. Line 378: This suggests that the completeness is wrong given that there are events missing above the completeness. Why is this?
- There are always missing events above the magnitude of completeness – only with synthetic tests or by comparison with other catalogs can we see it. I would explain it by the nature of noise – there are bursts that can cover the signal. Or there is coda of a stronger event or distant event. Or there might be two events nearby that either crash the detection method (mixture P and S from different events gives great error in location) or the later event is eliminated, because the detection methods usually have a blind window after a detection.
13. Line 383: How would the reduction of templates impact location accuracy (and hence focal mechanism accuracy)?
- This chapter is focused on detection (not location, or FM) results. We discuss the impact of reduced templates on detectability in terms of using single FM for all the templates (Fig 10a). If we reduce the templates to a single FM, there is no way to interpret FM.
 - If we would reduce the number of grid points but we keep the variation in FM, there will be an effect on location (since we assign the location to the closest template, not interpolate it). I did not run the simulations, but a theoretical example for our case: grid spacing increased from 1 km to 2 km reduces the number of templates 5.5 times and 7 of 9 events will be assigned to different templates. Therefore the 7 of 9 events would have the location perturbed by an extra ~ 1.1 km error.
 - An extra paragraph added to the section 4.3 Location by Synthetic Template Matching: If the number of template grid points is reduced but the variation in FM is kept, there will be an effect on location (since we assign the location to the closest template, not interpolate it). In our case, if the grid spacing is increased from 1 km to 2 km, it reduces the number of templates 5.5 times and 7 of 9 events will be assigned to different templates. Therefore, 7 of 9 events would have the location perturbed by an extra 1.1 km error on average.
14. Figure 8, panel c: I agree that this is clearly a real detection. It would help if the authors can state why this clear signal was not detected by PEPiN? This might help to demonstrate where the method outlined by the authors is most useful.
- In our previous study (Kaldy & Fischer, 2025), the efficiency of PePiN was tested on synthetic seismograms overlaid over real seismograms at a different local seismic network (GRSN - Litoměřice, Czech Republic). It shows that PePiN missed 18% of events above magnitude of completeness, mostly during the period of extensive noise, when the setting and geometry is similar to the one used in WEBNET for this study. Troubles were found when PEPiN was using borehole and surface sensors at similar locations; this issue does not apply in WEBNET.
 - For the real event in Fig 8c, the noise does not seem to be an issue at filter 1-10 Hz, but PePiN operates on a different frequency band (7-30 Hz). See the waveforms of that particular event in both frequency ranges below. On 7-30 Hz there is clearly a second signal colliding with the event visible also at 1-10 Hz. PePiN has likely recognized all

the arrivals and that has an issue with associating them to a single event, which resulted in high error in preliminary location and removing the detection.

- Explanation added to the text: PePiN missed the event (Fig. 8c) because in the frequency band 7-30 Hz, which PePiN uses for detection, there is an extra signal colliding with the event, causing trouble with phase association



In summary, I think this is a worthwhile and valuable contribution, and I am particularly interested in the ability of this method to estimate focal mechanisms. I suggest that the modest refinement to the text and some minor clarifications as outlined above will make this a more useful and understandable paper.

Recommendation: Revisions Required

Reviewer D:

This paper presents a new algorithm for event detection and focal mechanism identification using local synthetic seismograms coupled with a time-smearing cross-correlation aggregation procedure. The authors apply the new algorithm to the seismically active West Bohemia region, and compare it against two preexisting automated pick methods and the catalog for the area. The main novelty of the paper in this context is the ability to pick local seismicity with focal mechanisms in the absence of detailed reference templates or a high-quality 3D model.

Overall, the paper is thorough enough and presents its work well. I feel like there is a lack of motivation present in the introduction / discussion sections that could improve the readers' context and make the paper more impactful (see major comment below). In addition, there are a number of minor comments that I feel would improve the presentation and completeness of the paper. Consequently, I would recommend a moderate level of revision is required for this paper before publication.

Note: In general I do not review matters of English unless I feel like it may cause confusion in the understanding of the text's meaning. There is one case below that falls into that category; the authors should also do another pass for spelling and grammar errors through the paper.

Major Comments:

The overall motivation of the development of this algorithm needs to be elaborated on considerably. What are the operational / scientific motivations for performing this study? Is the intention to improve on past results, automate focal mechanism picking, both ...? While the final conclusion is that the new algorithm does not perform better than existing work in terms of event picks, while partially automating FM calculations, it is hard to tell what we have learned from this study overall. With clearer motivation would also come the opportunity to clearly assess the success of the project and potential directions for future improvements if warranted.

- The motivation for this study was to test a promising theoretical approach (synthetic templates for template matching) so others do not need to test it themselves.
- Modified first few lines of the introduction: While template matching is a highly sensitive method for detecting microseismicity, its traditional reliance on previously recorded earthquakes severely limits its utility in aseismic or newly monitored regions. This limitation poses a significant challenge for monitoring potentially induced seismicity associated with industrial operations—such as geothermal energy development (Zang et al., 2014), hydrocarbon extraction (Clarke et al., 2014; Grigoli et al., 2018), and CO₂ storage (White and Foxall, 2016) — which frequently occur in areas lacking a catalog of historical seismicity. Driven by recent advancements that allow for the efficient calculation of synthetic seismograms (Heimann et al., 2020, 2017), this study aims to bridge this methodological gap. Specifically, we investigate the extent to which a set of synthetic templates can reliably detect local seismicity, with the ultimate goal of establishing a robust monitoring framework for regions without prior known seismic activity.

Minor Comments:

73 - accuracy can only be improved; this statement is contingent on the appropriateness of the 3D model, and in low-data areas overfitting a 3D model to limited seismicity may result in poorer overall generalization of synthetic GFs to new locations. Please moderate and / or expand the discussion here.

- Paragraph modified: It is worth noting that if synthetic template matching proves valuable using a 1D velocity model, transitioning to a 3D velocity model offers a logical pathway for performance improvement. However, this is not a guaranteed enhancement. The success of a 3D approach is strictly contingent on the appropriateness of the model itself. In environments lacking dense data coverage, overfitting a 3D model to limited background seismicity can negatively impact the generalizability of synthetic Green's functions to new target locations, potentially making a well-calibrated 1D model the more stable choice.

74 - "number of templates, and merging of similarities from all stations"; the grammar of the list as written does not match the preceding "there is little difference...", in particular I can't understand the first clause. Is it supposed to read "selecting the number of templates"? Please adjust appropriately. It also appears to be missing the step of calculating the similarities between templates and observed data.

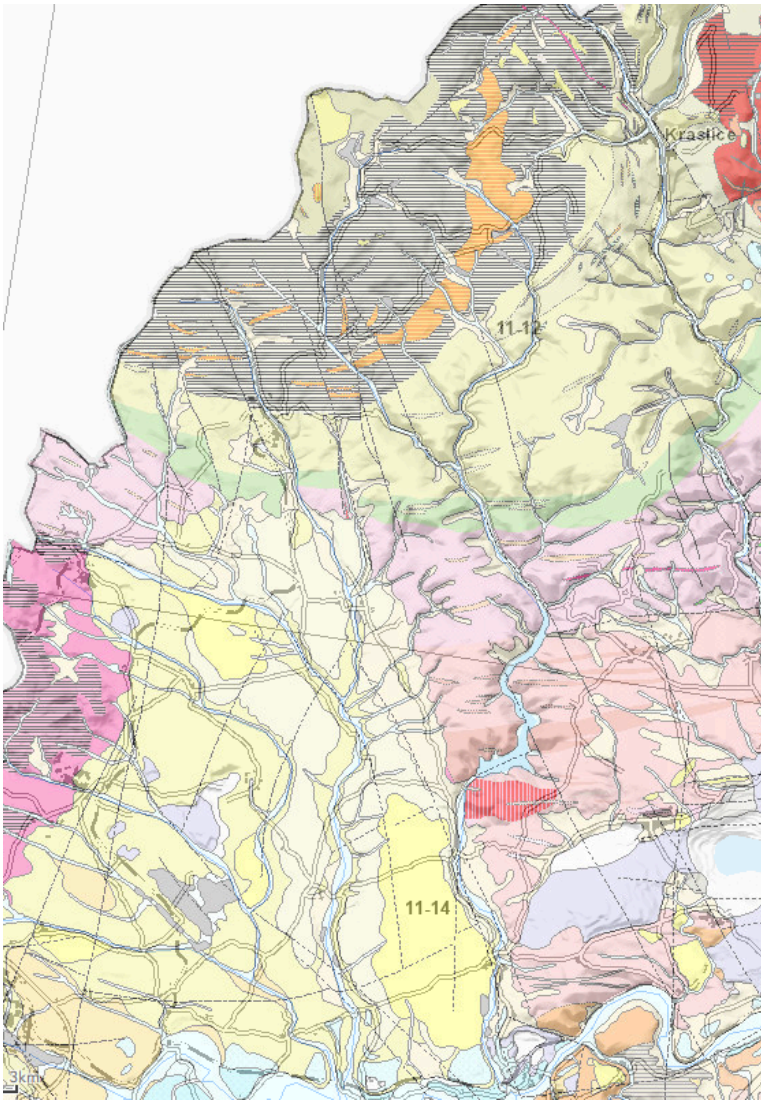
- Reformulated: There is a little difference in template matching procedure when using synthetic-waveform templates instead of the real-waveform ones. The difference lies in the number and spatial distribution of the templates and in the mechanism of merging the similarities:

83 - it would be helpful for readers to also convert the P-S time to an approximate distance for the appropriate velocity model.

- Modified: (hypocentral distance 10-20 km, $T_p - T_s \sim 1.7$ s)

Figure 1 - if it would be possible to reduce the clutter of minor basemap features from this map & overlay some geological context in addition to the human context, I feel it would be substantially improved.

- I am afraid the geological map (below) would raise more questions than it would improve the understanding.



143 - why should this computational limit be the goal? Both the generation of a large number of synthetic seismic templates & scanning data across them could be efficiently performed using a map-reduce paradigm within cloud computing platforms, probably using serverless + Docker for simplicity.

- The reasons for such are both practical and academic. If this is the first study on using synthetic templates for real detection, it is good to prove its usability on a small scale before going big. The practical reason is that a multicore desktop was what I had available in terms of computational power at the beginning of this study.
- But you are right, the next step is to bring the computation either serverless or to run it on our new computational PC (32 cores (64 threads), 128GB RAM, 12 GB GPU, 2TB system disk, 8 TB data disk). Since the process is already parallelized the second approach should be easy. This applies only if the process is run for a longer period of time (and extensive data load), not for real-time monitoring.
- Modified the paragraph: The goal is to create a grid of synthetic templates that detects all earthquakes in the area of interest, regardless of their exact location or focal mechanism. A dense grid with a great variation of focal mechanisms would serve the detection well, but it significantly increases the computational burden. While cloud computing enables the use of highly dense grids, our objective is to provide a highly accessible, rapid-deployment tool that can be executed on a standard multi-core computer for on-site, near-real-time monitoring where cloud infrastructure may be impractical or unavailable. To achieve this efficiently, tools

such as Pyrocko (Heimann et al., 2017, 2019, 2020) facilitate the generation of synthetic waveforms using established velocity models, and a modified version of ObsPy (Beyreuther et al., 2010) handles the template matching detection.

Figure 2 - please annotate each subplot with the normalised cross-correlation for that trace pair.

- I am sorry no cross-correlation was calculated for this pair. This display is here only to express the visible differences between real and synthetic if the synthetic represents similar location and focal mechanism. It presents the uncertainty and incompleteness of the 1D velocity model and synthetic seismograms in total.

239 - please clarify whether the time smearing is applied to the P and S arrivals independently or if the same correction is applied to both

- Whole paragraph modified, also based on the opinion of the second reviewer: A key distinction between template matching using real versus synthetic templates does not lie in the morphology of the P- and S-waveforms themselves, but in the relative timing of their arrivals across the station network. Synthetic templates derived from a 1D velocity model inherently contain inaccuracies in predicted travel times due to model simplifications, lateral velocity heterogeneity, and potential anisotropy. As a result, while real and synthetic P- and S-wave arrivals may coincide at some stations, they can exhibit measurable time shifts at others (e.g., LBC and POC in Fig. 2).

In regions lacking prior seismicity, these time offsets cannot be directly constrained from observations. Instead, they must be inferred from the uncertainty of the velocity model. In our case, we estimate the expected error in relative arrival times across stations as σ_{terr} to ~ 0.43 s.

For more complex synthetic waveforms, misalignment between real and synthetic arrivals could be mitigated by defining separate matched-filter windows for P and S phases. However, in the context of induced seismicity at local scales—which is the focus of this study—both real and synthetic waveforms are relatively simple in shape. Under these conditions, using separate P- and S-wave windows substantially increases the rate of false detections. To minimize such false triggers, we therefore adopt a single combined P+S matched-filter window.

Figure 4 - this figure would be more intuitive if the synthetic and real data were more closely aligned

- There is quite a difference in relative wave arrivals since the real and synthetic waveforms are for earthquakes separated by 5 km, overlaying would then work only for a single station...

271 - this procedure for computing the focal mechanism should be put into mathematics within the paper

- A formula explaining the creation of the moment tensor assigned to the detection is put into the paragraph, which is modified accordingly.

The focal mechanism FM_{int} assigned to a detected event is interpreted from N FM_k of templates detecting the event and having a hypocenter similar to that of the detecting template with the highest similarity. The interpreted FM_{int} is a linear combination the FM_k (Fig. 3 left) converted to moment tensors M_k , weighted by its similarity s_k , reduced by the minimal similarity $s_{min} = \min s_k$, and the resulting moment tensor M_{res} converted back to FM_{int} : $M_{res} = \sum_{k=1}^N (s_k - s_{min}) M_k$. The conversion between FM and moment tensor is performed by Pyrocko's *moment_tensor* function (Heimann et al., 2017).

Listing 1 - this listing is incomplete (uses non-imported functions); in addition the function arguments in the second function are uninterpretable without the context. Personally, I would recommend removing the listing in favor of the code release alone as I don't feel it adds to the paper.

- Removed, text modified accordingly.

Figure 9 - if possible, please adjust the plot so that the NS projection is horizontal and the depth projection is vertical as the current presentation is unintuitive (recognizing that it is currently arranged so that north is the same for both subplots; potentially rotating the entire plot 90 degrees and adding a north arrow is an option which at least to me makes more sense after checking by rotating my screen!). The use of the same box notation in the NS-EW and NS-Z subplots is also confusing (since the box is roughly the same size in both plots and uses the same symbology, it intuitively looks like the same thing which makes the RHS plot look like a map projection rather than a cross section)

- A similar display is in Fig. 6.
- I am afraid I would be facing the same understanding issues if the display was as you suggest.
- If the point was to understand the structure of the fault (seismicity location along fault), I would make the cross sections parallel and perpendicular to the fault plane (and then the depth would be pointing down). But here we aim to express location errors.
- To make it more clear to people who are used to different displays, a comment “Map & North-South depth cross-section.” is added to the beginning of the captions (both Fig 6 and 9)
- Help for looking at this display “Map & North-South depth cross-section”:
 - It can be imagined as a box with all the events inside. You get the map view when looking from the top and the depth displays when looking from the sides. To plot the box on paper you “open” the North-South correction to the right and East-West cross section to below (we do not display this here).

423 - I'm confused here whether the focal mechanisms are those derived by the template matching algorithm or the manually interpreted ones.

- Here we refer to a location which comes in the same catalog as manual FM.
- Modified text: the location for FM (Vavryčuk et al., 2022), which uses 22 - 23 stations.

Recommendation: Revisions Required
