Supplementary Material for
Dense Seismic Array Monitoring of the Glacier Tongue of Isunnguata Sermia, West Greenland

Nicolas Paris *[1], **Florent Gimbert**[1], **Tifenn Le Bris**[1], **Stephen J. Livingstone**[2], **Samuel H. Doyle**[2,3], **Alexandre Michel**[1], **Andrew J. Sole**[2], **Albanne Lecointre**[4], **Laura Pinzon-Rincon**[5], **Gregor Hillers**[6], **Roméo Courbis**[6], **Philippe Roux**[4], **Guilhem Barruol**[1], **Elizabeth A. Bagshaw**[7], **Thomas R. Chudley**[7,8], **Lisa Craw**[9], **Laura A. Edwards**[10], **Adrien Gilbert**[1], **Jonathan D. Hawkins**[9], **Ryan N. Ing**[11], **Andrew H. Jones**[12], **Angus Moffat**[2], **Matthew W. Peacey**[3,7], **Michael R. Prior-Jones**[9], **Neil Ross**[13], **Arnaud Reboud**[1], **Robert D. Storrar**[14], **Sian C. Thorpe**[2], **Remy Veness**[12], **Tun Jan Young**[15]

[1]Université Grenoble Alpes, CNRS, INRAE, IRD, Grenoble INP, IGE, Grenoble, France, [2]School of Geography and Planning, University of Sheffield, Sheffield, UK, [3]Centre for Glaciology, Department of Geography and Earth Sciences, Aberystwyth University, Aberystwyth, SY23 3DB, UK, [4]ISTerre, Université Grenoble Alpes, CNRS, Université Savoie Mont Blanc, IRD, Université Gustave Eiffel, Grenoble, France, [5]German Research Centre for Geosciences (GFZ), Potsdam, Germany, [6]Institute of Seismology, Department of Geosciences and Geography, University of Helsinki, Helsinki, Finland, [7]School of Geographical Sciences, University of Bristol, University Road, Bristol BS8 1SS, UK, [8]Department of Geography, Durham University, Durham, UK, [9]School of Earth and Environmental Sciences, Cardiff University, Cardiff, GB, [10]School of Biological and Environmental Sciences, Liverpool John Moores University, Liverpool, UK, [11]School of Geosciences, University of Edinburgh, Edinburgh, UK, [12]Institute of Social Sciences: Geography, Environment and Planning, Sheffield Hallam University, Sheffield, UK, [13]School of Geography, Politics and Sociology, Newcastle University, Newcastle upon Tyne, UK, [14]Department of the Natural and Built Environment, Sheffield Hallam University, Sheffield, UK, [15]School of Geography and Sustainable Development, University of St Andrews, St Andrews, UK

S1 Hourly noise correlation functions calculation

Before computing the hourly correlations, we bandpass-filtered the waveforms between 0.1 and 100.0 Hz and clipped amplitudes at five times the standard deviation to suppress short-duration, high-amplitude noise. We then applied spectral whitening, followed by 1-bit normalization. Then, we rotated the North and East components into radial and transverse components based on the azimuth of each station pair. For a given station-pair, we computed correlations using 10-minute windows with an 8-minute overlap, yielding 26 correlation windows per hour. After normalizing all correlations by their maximum amplitude, we averaged the 26 normalized correlations to produce an hourly correlation. We repeated this process for each hour, station-pair, and component combination.

S2 GNSS desynchronization: estimation of clock drift coefficients

We first examined the GNSS logs of all nodes to identify the stations that were most likely to have remained synchronized throughout the recording period. Five of these stations were selected for subsequent processing and are hereafter referred to as reference stations. We then inspected the ZZ correlograms between each reference station and all other stations to detect periods of desynchronization and to manually define the start and end times of each desynchronization episode.

For each station that experienced a desynchronization episode, we estimated the clock drift coefficient and time offset relative to reference stations that were assumed to remain synchronized throughout the deployment. These parameters were inferred from the temporal alignment of ambient noise cross-correlation functions under the assumption that maximizing the energy of the stacked correlograms corresponds to optimal time synchronization. The following procedure describes the estimation of the desynchronization parameters of a station:

- 1. Select any station, referred to as station A.
- 2. Choose a reference station and create the ZZ correlogram between station A and the reference station.
- 3. Perform a grid search on the drift coefficient and offset (only for 2024 nodes), where, for each tested combination of parameters, shift the waveforms to account for the expected shift.

*Corresponding author: nicolas.paris3@univ-grenoble-alpes.fr

- 4. Stack all hourly correlations and compute the energy of the stacked waveform to find the pair of parameters that leads to the largest energy.
- 5. Repeat steps 2–4, choosing a different reference station each time. This leads to 5 different estimates of each parameter for station A.
- 6. Average the 5 estimates of each parameter to obtain the estimated drift coefficient and offset of station A.

The estimation of the desynchronization parameters for node 1097 in summer 2024 required a different approach. Because this node was desynchronized throughout the entire recording period, its desynchronization offset could not be determined using the same method as for the other stations. We therefore first estimated only its drift coefficient, following the same approach as with other nodes, but constraining the offset to zero. After correcting the node's waveforms for the drift-induced time shift, we cross-correlated them with a co-located DAS channel during the sequence of sledgehammer shots performed next to station 1097. Given that the DAS logs indicate that its GNSS synchronization was valid at the time of the shots (Figure S6), and considering that the expected drift is only on the order of a few milliseconds per day, the DAS data is assumed to be correctly synchronized. We therefore estimated the remaining time lag required to align the signals from the maximum of the cross-correlation to estimate the offset of node 1097 (Figure S7).

S3 Matched Field Processing computation

We used Matched Field Processing to create event catalogs with the dense seismic node arrays deployed in 2023 and 2024. Event locations and phase velocities were optimized using the Nelder–Mead algorithm (Nelder and Mead, 1965). Because this algorithm performs local rather than global optimization, we evaluated multiple starting points to reduce the risk of converging to a local minimum of the cost function. To determine an efficient trade-off between accuracy and computational cost, we first ran synthetic tests to identify the optimal number of starting points. We then used this configuration to detect and locate the 2023 and 2024 events.

S3.1 Synthetic test: Determination of the optimum starting point number

To determine the optimal number of starting points we conducted synthetic tests in which noise was added to the synthetic waveforms. This allowed us to verify that the chosen number of starting points remains sufficient even for events with low optimum Bartlett output (as low as 0.3-–0.4). For these tests, we generated vertical-component synthetic seismograms using a Ricker wavelet with a central frequency of 5 Hz. The waveforms were time-shifted according to the source–station distance, assuming a phase velocity of 1650 m/s. To reduce the Bartlett output and produce events with optimal values in the 0.3-–0.4 range, we added white noise with a normal distribution whose power increases linearly with source–station distance.

We constructed a grid of surface events and, for each event, optimized the source location and phase velocity using 300 starting points. These starting points were uniformly distributed in 3-D, with depths ranging from 0 to 650 m. To evaluate how many starting points are required, we formed subsets of size N (from 1 to 300) and performed the analysis using only those N starting points. For each subset, we computed the Euclidean distance between the optimized and true source locations and retained the smallest distance as the optimization error. This procedure yields, for each event, a measurement of the location error as a function of the number of starting points N used in the optimization. Then, to determine the optimal number of starting points across all events, we calculated the median location error as a function of N (Figure S9c). We then evaluated the rate at which the error decreases with increasing N (Figure S9d) and applied an arbitrary threshold of 0.1%. Based on this criterion, we decided to use 125 starting points for the 2023 dataset and 150 for the 2024 dataset.

However, with the 2024 array geometry, a region of large errors persists (Figure S9b) due to the higher station density and the resulting increase in array-response complexity. To improve convergence in this area, we added 100 additional starting points uniformly distributed within the high-error region (Figure S9e), bringing the total number of starting points for the 2024 array geometry to 250.

S3.2 Applying MFP to 2023 and 2024 data

We located the seismic events using Matched Field Processing applied to the vertical component of the seismic nodes, after removing stations affected by significant tilt and correcting for clock desynchronization. The calculations were performed in non-overlapping 2-s time windows in 15 frequency bands with central frequencies ranging from 0.4 to 16 Hz, with a 0.1 Hz frequency discretization. For each window, we initialized the optimization with 125 starting points for the 2023 array and 250 for the 2024 array. An initial velocity of 1800 m/s was prescribed in all cases to assess convergence toward the expected Rayleigh wave phase velocity. In our analysis we focused on the 4–6 Hz frequency band, discretized in 21 frequencies, and selected the optimized location with the highest Bartlett output as the source location in each time window.

S4 Synthetic tests

All synthetic tests used a Ricker wavelet on the vertical component, with a central frequency of 5 Hz, time-shifted according to the source–station distance with a constant 1650 m/s phase velocity.

S4.1 Resolution on event location

The resolution tests were performed on a grid of synthetic events using the same parameterization as in the determination of the optimal number of starting points (Section S3.1), but with varying noise levels to simulate different maximum Bartlett outputs. Each synthetic event was generated 30 times with independent noise realizations. For each realization, we used the same set of starting points as in the real-data analysis (Section S3.2) and retained the optimized location with the highest Bartlett output as the solution. We then computed the horizontal and vertical distances between the true source location and the corresponding optimized locations across the 30 realizations and calculated the average errors. This procedure was repeated for all source locations, yielding maps of average horizontal and vertical error.

S4.2 Parameters affecting Bartlett output: Tilt

We assessed the impact of station tilt by generating 15,000 synthetic event realizations "recorded" by randomly tilted stations. Due to the tilt, the signal on the vertical component is no longer purely vertical. To account for this, we generated a 90° phase-shifted version of the Ricker wavelet and applied it to the radial component to simulate Rayleigh waves, which dominate the signal in the 4—6 Hz band. For each realization, we applied a rotation to the signal to mimic the effect of tilt, using a random tilt direction between 0 and 360° and random tilt angles between 0° and 30°. We then computed the Bartlett output at the true source location using the tilted vertical component waveforms.

S4.3 Parameters affecting Bartlett output: Noise level

To evaluate the effect of noise level on the maximum Bartlett output, we generated synthetic events with varying Bartlett outputs using the noise level from spring 2023, performing an optimization with the starting point set at the true source location. The SNR was then adjusted to reflect the noise levels of 2024, under the assumption that the signal amplitude remained unchanged between spring 2023 and 2024. After this SNR correction, a second MFP was performed, again starting at the true source location, to obtain the optimized Bartlett output with the updated noise level.

References

Nelder, J. A. and Mead, R. A Simplex Method for Function Minimization. *The Computer Journal*, 7(4):308–313, 01 1965. doi: 10.1093/comjnl/7.4.308.

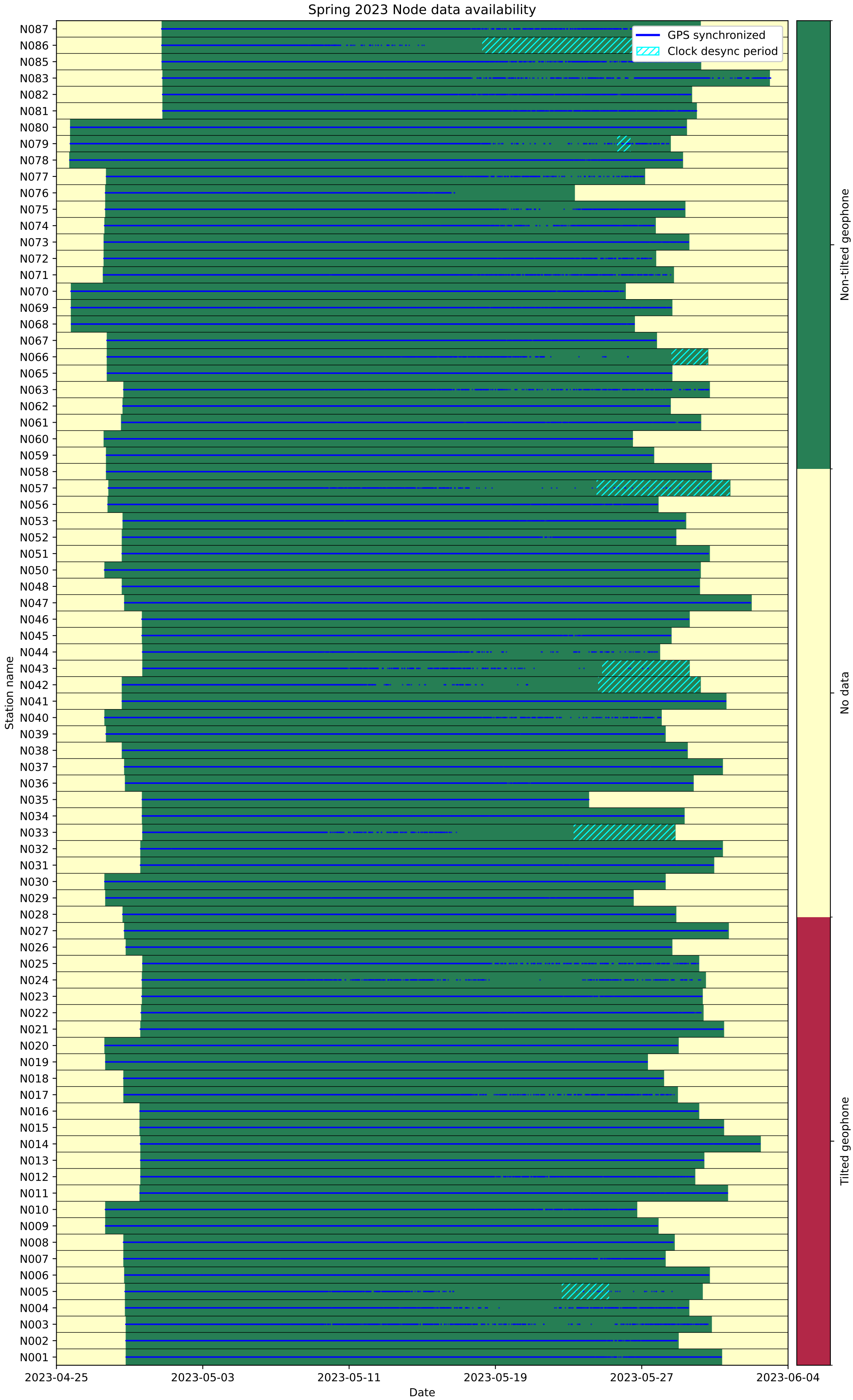

Figure S1 Data availability of the spring 2023 dataset. The colormap indicates data availability and quality on the vertical component. Blue dots indicate the when the unit received GNSS signal. Hashed cyan areas indicate the time periods where we observed time desynchronization in ambient noise correlation functions.

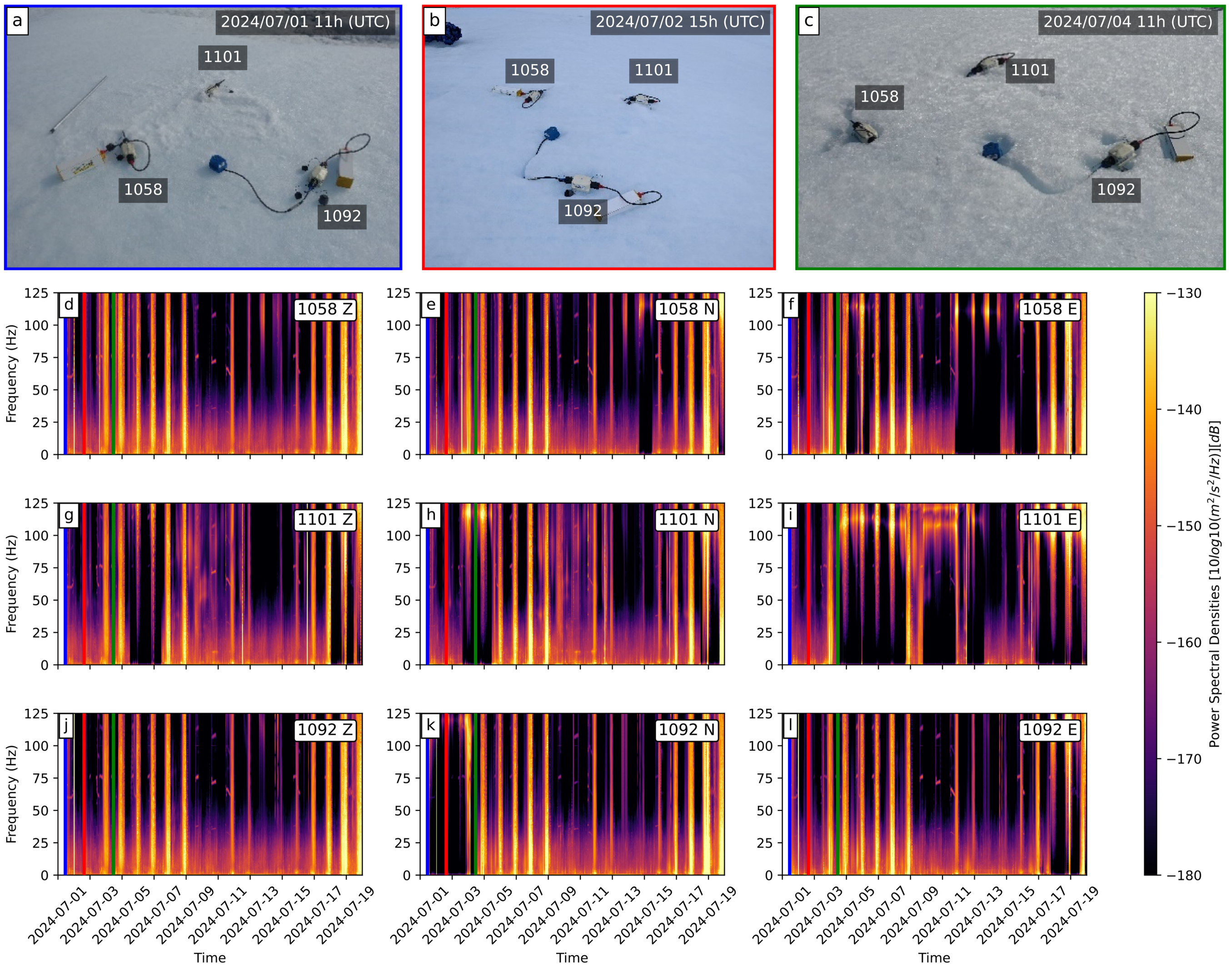

Figure S2 a–c) Photos of the co-located test nodes 1092, 1101, and 1058 at different dates. d–l) Power Spectral Density of the test nodes on all three components (Z, N, E). The vertical lines mark timestamps of photos in panels a–c), with the their color matching that of the photo contours.

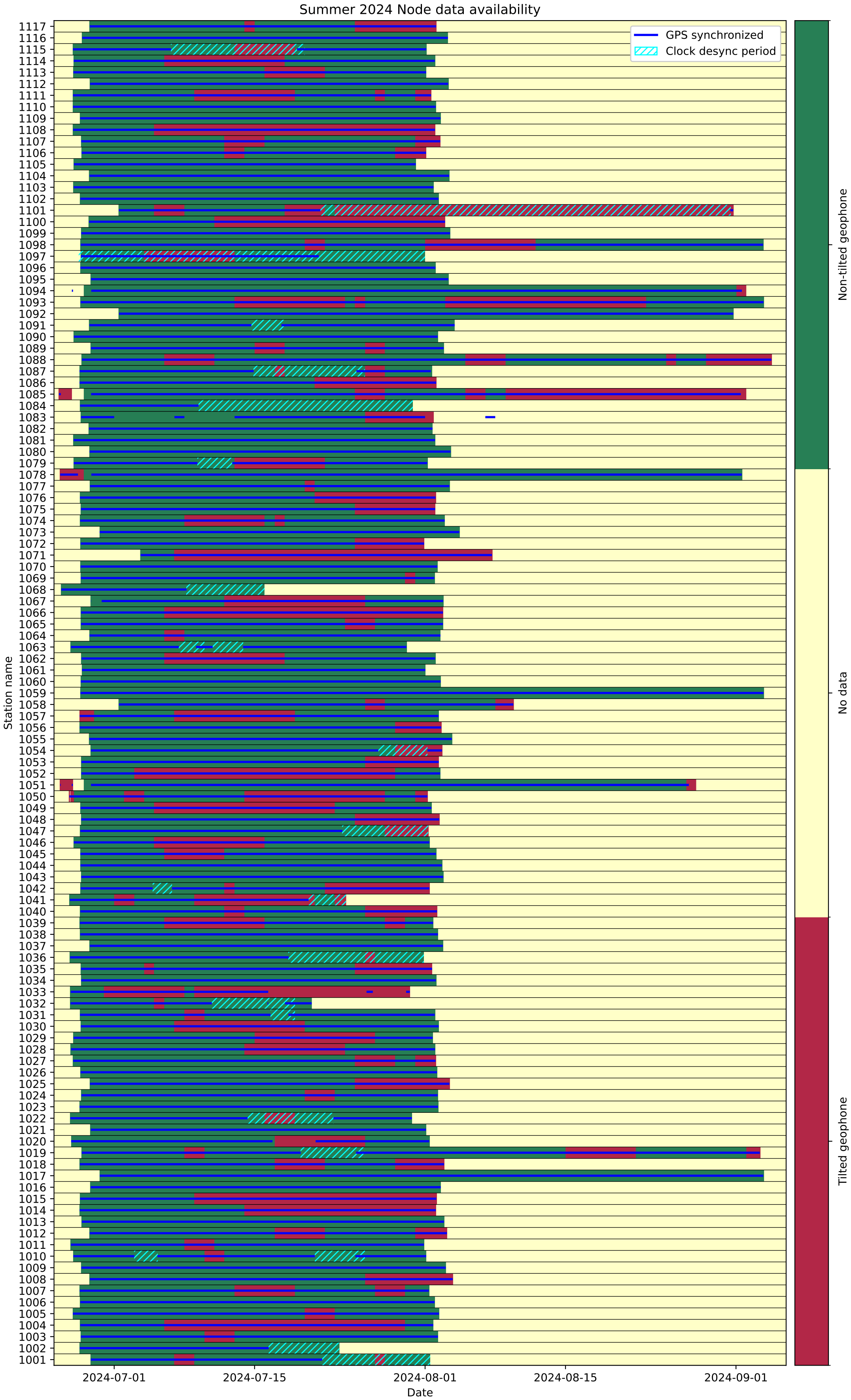

Figure S3 Same as Figure S1 but for the summer 2024 nodes.

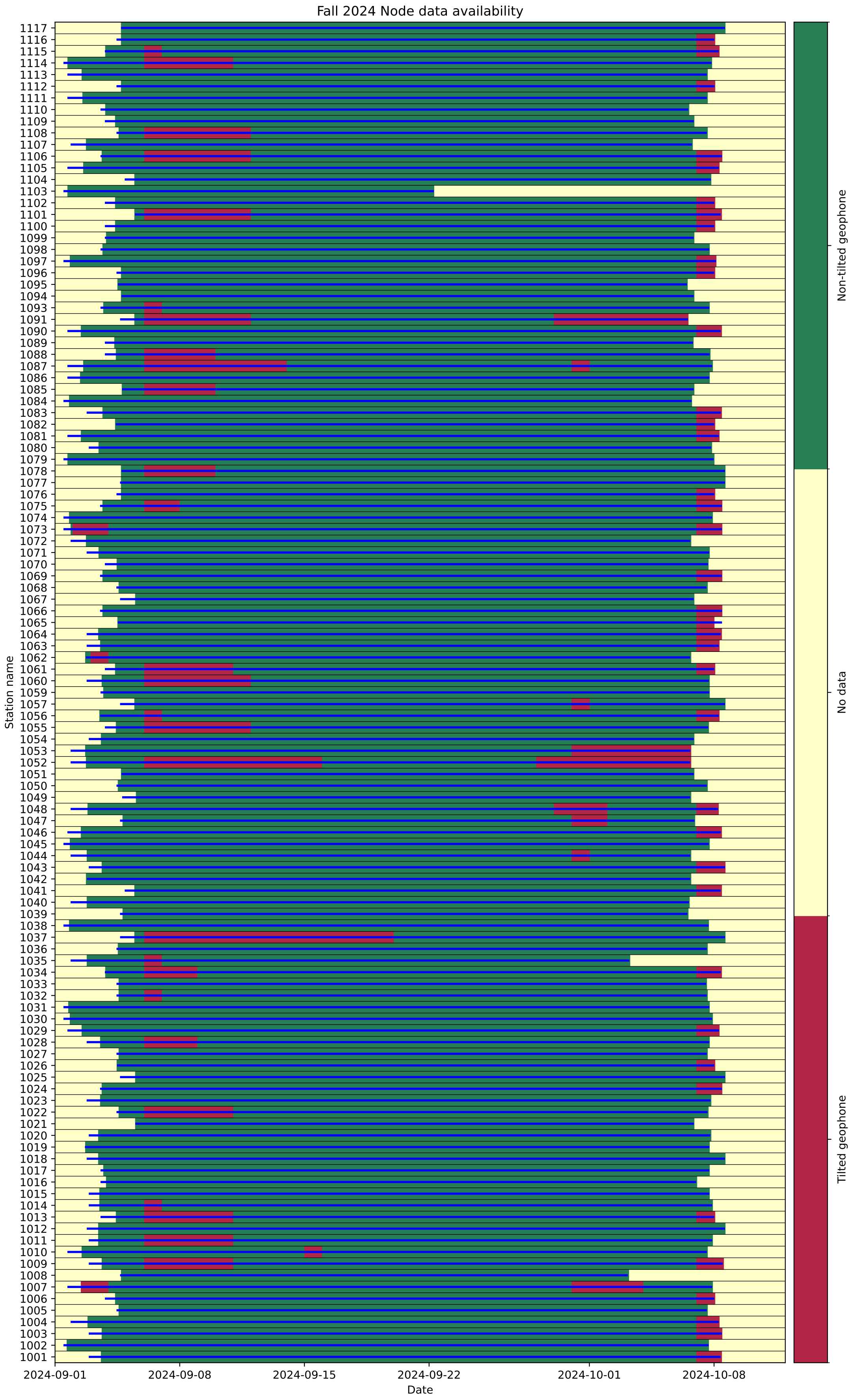

Figure S4 Same as Figure S1 but for the fall 2024 nodes.

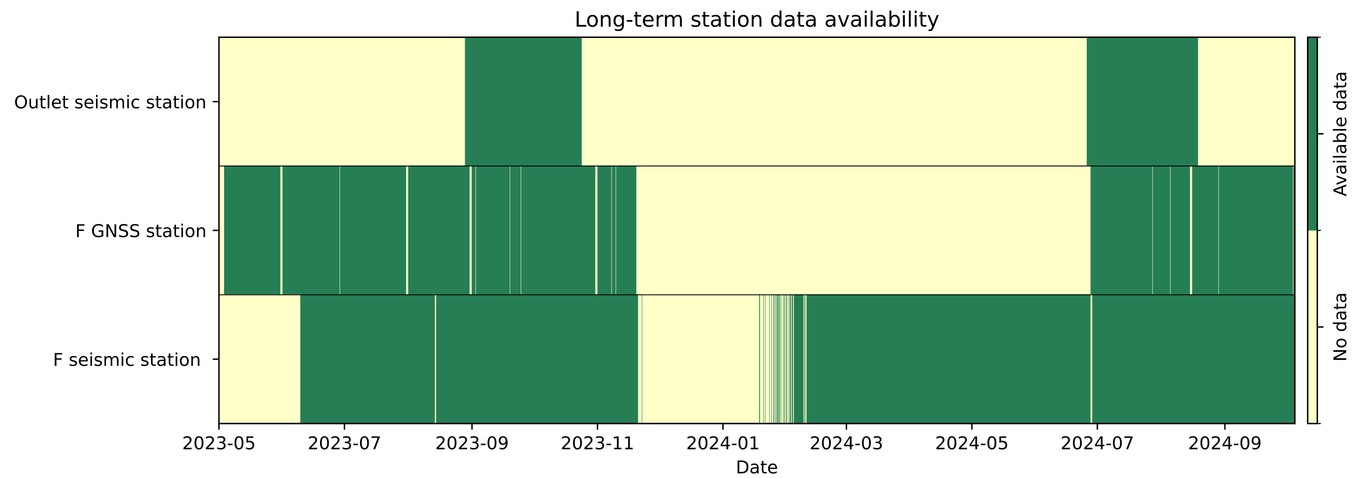

Figure S5 Data availability of long-term deployments.

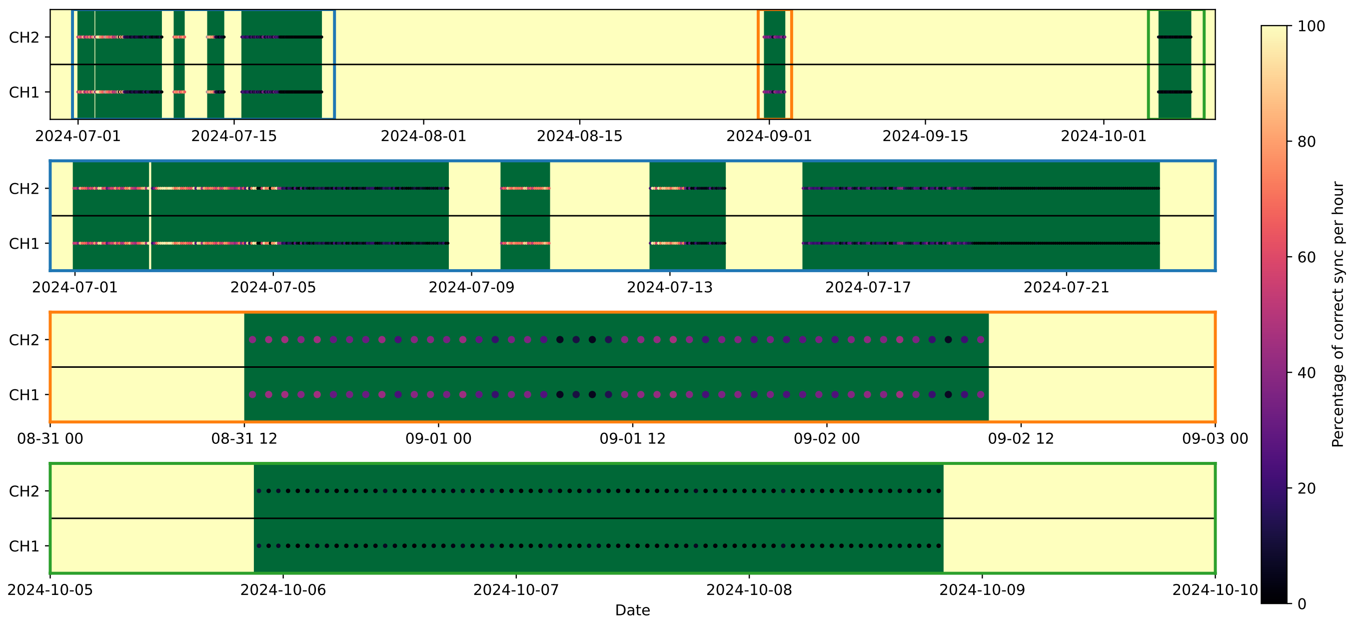

Figure S6 DAS data availability. The background green and yellow colors indicate times when there is and is not available data. The color of the dots indicate the percentage of time when the logs indicate the data is correctly synchronized within each hour.

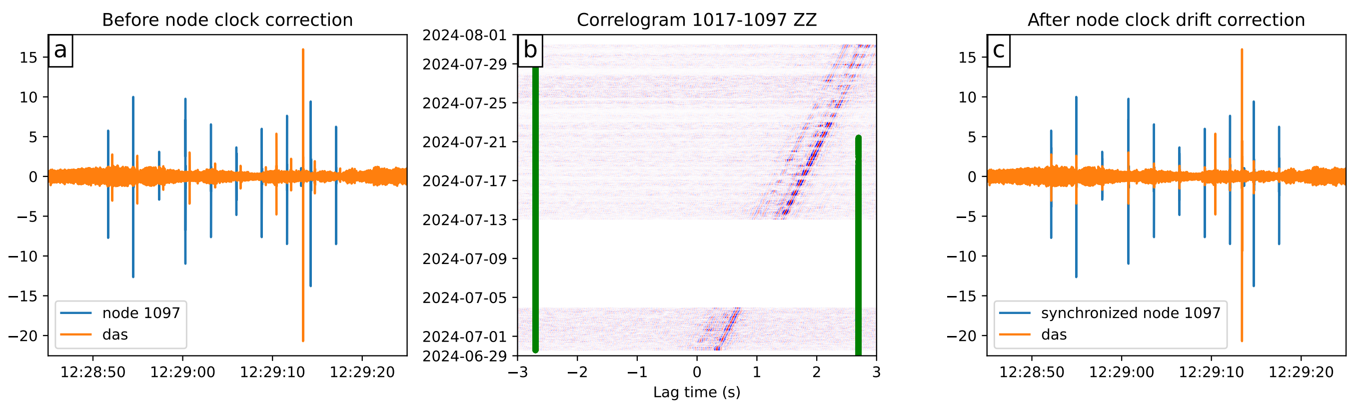

Figure S7 a) Comparison between the signal recorded by node 1097 along its E-W component and its co-located DAS channel during a series of sledgehammer blows nearby. b) ZZ Correlogram between nodes 1017 and 1097. The green lines on the left and right indicate when nodes 1017 and 1097 receive GNSS signal, respectively. Same as panel a), but after correcting for the desynchronization of node 1097.

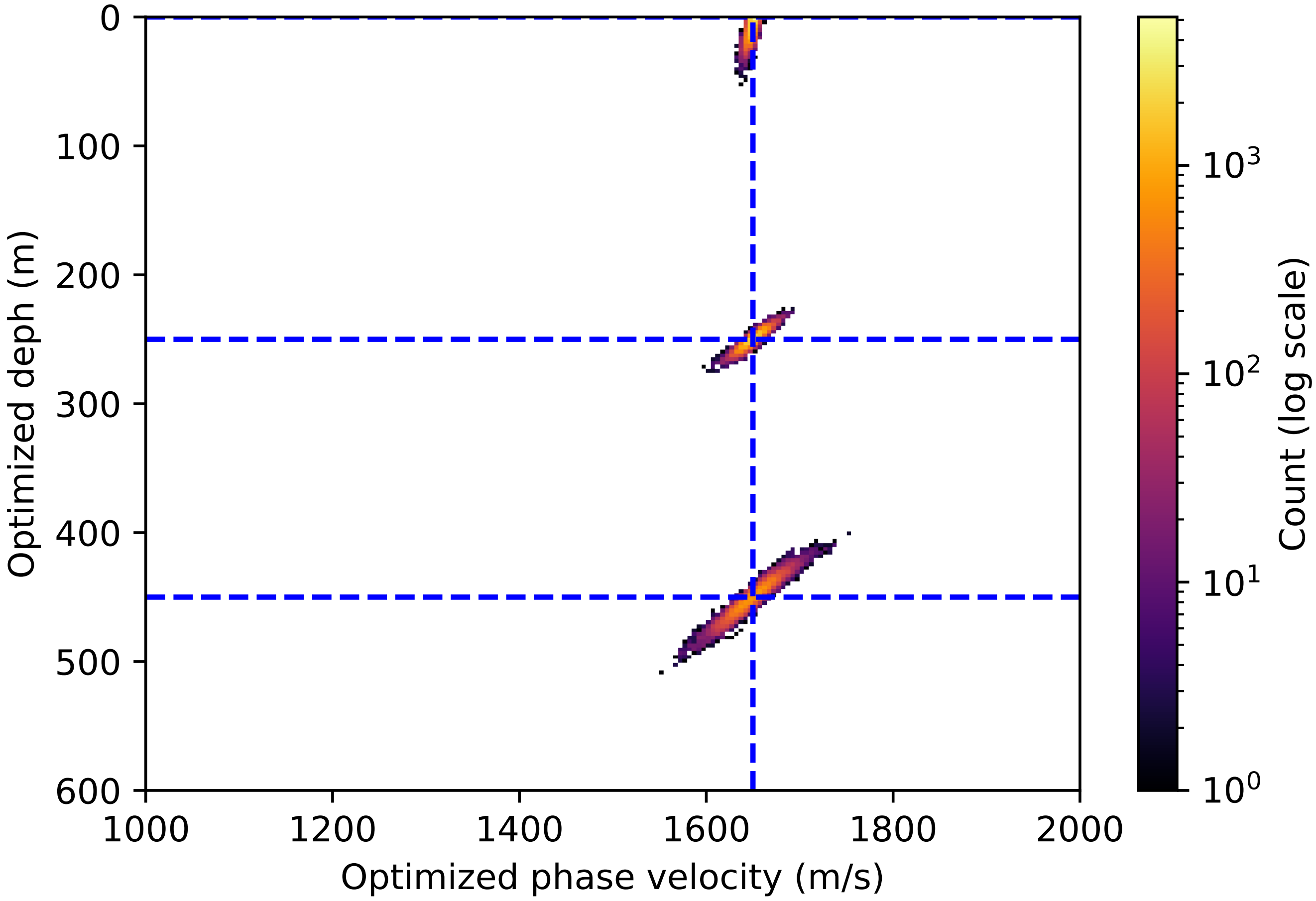

Figure S8 Optimized velocity vs optimized depth from high BO synthetic events at 0 m, 250 m and 450 m depth in the spring 2023 array configuration. Blue dashed lines indicate the true depth and velocity of the events. The deeper the event, the larger the spread in optimized phase velocity, which correlates negatively with optimized depth.

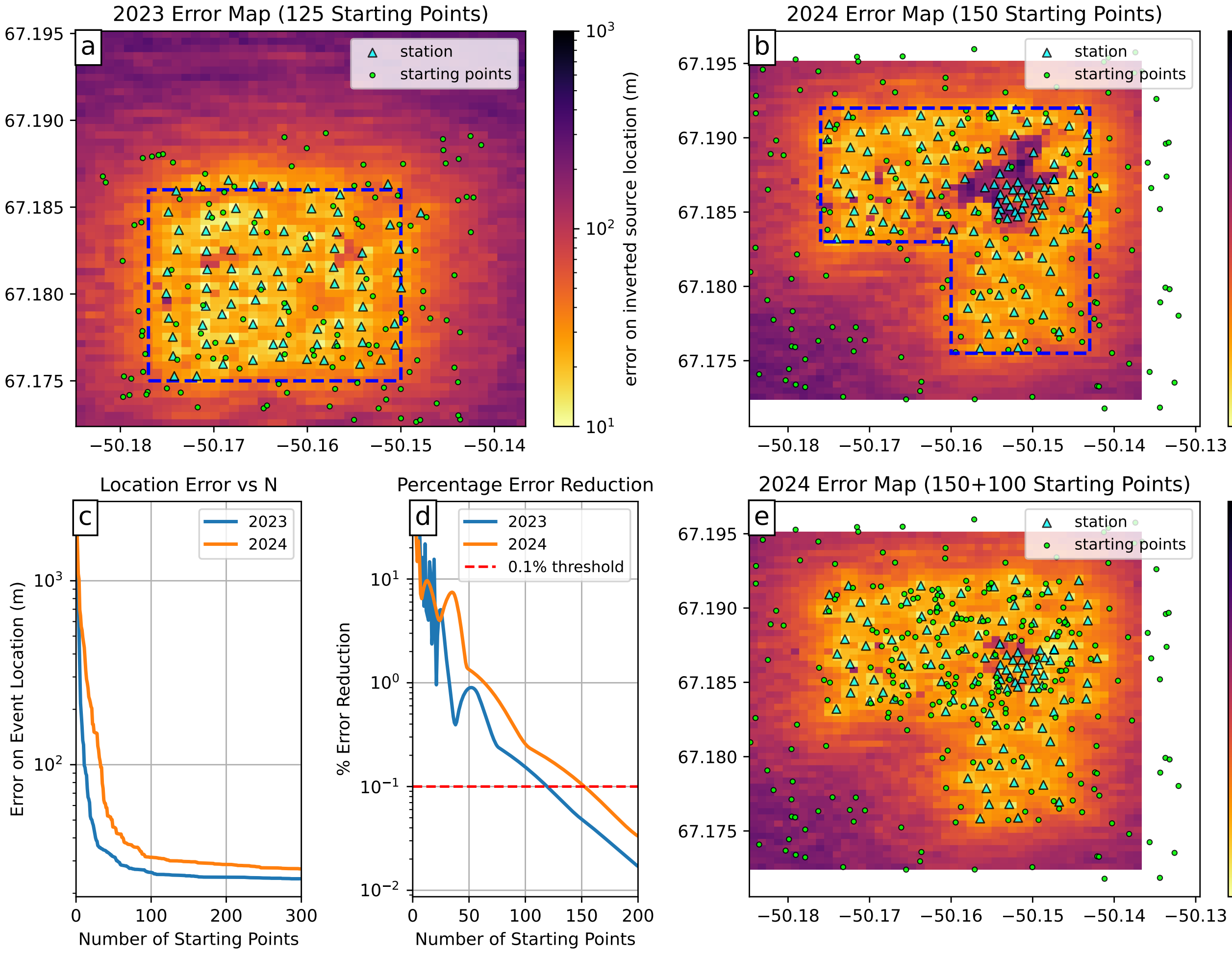

Figure S9 Averaged error on the inverted location for a) 2023 and b) 2024 arrays, using 125 and 150 starting points, respectively. The dots indicate the location of the starting points and the triangles indicate the station location. The blue dashed polygon indicate the event location used in panel c). c) Average error on the location as a function of the number of starting points for the 2023 and 2024 arrays. d) Relative location error decrease between successive number of starting points for the 2023 and 2024 array. The red dashed line indicates the 0.1% threshold used to choose the number of starting points for the inversion. e) same as panel b) but with 100 additional starting points in the northern part of the array.