SUPPORTING ON-LINE MATERIALS FOR: Weighing geophysical data with trans-dimensional algorithms: An earthquake location case study

Nicola Piana Agostinetti,[1,2], Alberto Malinverno,[3], Thomas Bodin[4], Christina Dahner[5], Savka Dineva[5,6]and Eduard Kissling[7]

[1]Department of Earth and Environmental Sciences, Universitaḋi Milano Bicocca, Piazza della Scienza 1, I-20126 Milano, Italy
[2]Department of Geology, University of Vienna, Vienna, Austria
[3]Lamont-Doherty Earth Observatory, Columbia University, NY, USA
[4]Univ Lyon, Univ Lyon 1, ENSL, CNRS, LGL-TPE, F-69622, Villeurbanne, France
[5]Luossavaara-Kiirunavaara AB, Kiruna, Sweden
[6]Department of Civil, Environmental and Natural Resources Engineering, Lulea, Sweden
[7]Department of Earth Sciences, ETH Zurich, Switzerland

Corresponding author: Nicola Piana Agostinetti, `nicola.pianaagostinetti@unimib.it`

Abstract

In geophysical inverse problems, the distribution of physical properties in an Earth model is inferred from a set of measured data. A necessary step is to select data that are best suited to the problem at hand. This step is performed ahead of solving the inverse problem, generally on the basis of expert knowledge. However, expert-opinion can introduce bias based on pre-conceptions. Here we apply a trans-dimensional algorithm to automatically weigh data on the basis of how consistent they are with the fundamental assumptions made to solve the inverse problem. We demonstrate this approach by inverting arrival times for the location of a seismic source in an elastic half space, under the assumptions of a point source and constant velocities. The key advantage is that the data do no longer need to be selected by an expert, but they are assigned varying weights during the inversion procedure.

1 Materials and Methods

1.1 Azimuthal coverage.

Azimuthal coverage of seismic sensors is a key parameter that affects the reliability of seismic event locations. In crustal studies, seismic sensors are typically located at the surface, i.e., on a single plane that contains the epicenter. In this case, the "azimuthal gap" denotes the greatest angle between the epicenter and any two seismic sensors that record the seismic waves propagating from the event. In general, a large azimuthal gap is associated to large location uncertainties, as it happens for example for offshore seismic events when all seismic sensors are located on shore. The same azimuthal gap definition does not hold for a 3D distribution of seismic sensors. In our study, we compute the azimuthal gap of a 3D distribution of seismic sensors as follows.

1. We project all seismic sensors onto a sphere centered in the event source location (like stars in the sky), so that the position of each sensor has two angular coordinates, azimuth ϕ and altitude λ;
2. For each sensor on the sphere, we compute the minimum angular distance to every other sensor on the sphere;
3. The "azimuthal gap" of the sensor configuration is the largest of these minimum angular distances.

This definition of azimuthal gap quantifies how sparse is the constellation of seismic sensors around a given event location and can be related to event location uncertainties. This definition also applies to the case where sensors are distributed on a plane at the surface that contains the event epicenter.

1.2 Experimental Design.

In our study, we performed three different experiments based on two data sets of seismic measurements. Raw seismic waveforms were recorded in the Kiirunavaara mine (Sweden) by a seismic network (253 seismic sensors spanning the 3D volume around the ore body (Dineva et al., 2022)). From the raw seismic waveforms, P- and S-wave arrival times were automatically retrieved and manually revised by the local seismic system provider (Institute of Mine Seismology, IMS). These arrival times are the data used to solve the geophysical inverse problem of locating the seismic source. In the first data set, the seismic source is a man-made blast used to calibrate the seismic system. The position of the seismic source is known with an accuracy less than 1 meter. This data set can be used as the input of a robust field test, as the source location is known with an accuracy higher than the half-width of the characteristic wavelength of the P-wave (given a blast corner frequency close to 1000 Hz, if we assume Vp = 6600 m/s, for example, the half-wavelength is 3.3 m.). Due to the limited amount of explosive used (5 kg), seismic waves have been clearly recorded at 57 seismic sensors only (with 57 P- and 9 S-wave arrival times), which translates in a minimum and maximum sensor distance from the source of 83 and 710 meters, respectively. The second data set is based on P- and S-waves generated during a M_w 4.2 event that occurred on May 18th, 2020. This large event partially destroyed the mine infrastructure in a section about 1300 meters wide(Dineva et al., 2022). The resulting P- and S-waves were recorded at 151 seismic sensors (151 P-wave and 81 S-wave arrival times) as far as about 2130 meters from the preliminary seismic source location. Our preliminary seismic source is the official location (Dineva et al., 2022), even though the damaged area may be as large as 700×250 meters. For the present study, we specifically revised all P- and S-wave arrival times associated to this event to include as many seismic sensors as possible. It is worth noting that some seismic sensors close to the rupture area were destroyed during the event, limiting the availability of data near the source. The closest sensor to the preliminary seismic source is about 140 meters away.

1.3 The reference solution: McMC location of a seismic event

Our fist experiment obtains a reference solution that consists of estimated source locations obtained using only sensors within a maximum distance from the calibration blast (which is equivalent to have equal weight for all stations within the max distance and ignore stations that are beyond the maximum distance). By considering a range of possible maximum sensor distances, we aim to reproduce the results that would be obtained by different expert opinions.

The observed arrival time data in a vector $\mathbf{d}$ are $t^{\text{obs}}_{x,i}$, where $x = P$ or S, $i = 1, \dots, N$, and $N = N_P + N_S = 66$, with N_P and N_S being number of P- and S-wave arrival times, respectively. We apply a standard Markov chain Monte Carlo approach to solve the inverse problem of locating the source. We make four simplifying assumptions: (1) The rock volume is a homogeneous half-space; (2) The seismic event is a point source; (3) The arrival times have associated uncertainties equal to the sampling rate (i.e., $\sigma_0 = 1/6000$ s); and (4) The covariance matrix of the data errors $\mathbf{C}^*_e$ is diagonal, i.e. $\mathbf{C}^*_e = \sigma_0^2\,\mathbf{I}$, with $\mathbf{I}$ being the identity matrix (see (Riva & Piana Agostinetti, 2023) for details on the methodology).

The model vector $\mathbf{m}$ in the inverse problem contains eight parameters. Six parameters are related to the physical model: the coordinates of the seismic source (X_s, Y_s, Z_s), the origin time measured with respect to the first P-wave arrival time (OT), the P-wave velocity of the half-space (V_P), and the ratio between P-wave velocity and S-wave velocity in the half-space (V_P/V_S). Two additional "hyperparameters" (Malinverno & Briggs, 2004) π_P and π_S, defined below, control the data uncertainties. Thus, $\mathbf{m} = (X_s, Y_s, Z_s, OT_s, V_P, V_P/V_S, \pi_P, \pi_S)$. In our experiment, all prior probability distributions are uniform within the minimum and maximum values given in Table 1.

For the given assumptions, the seismic ray paths for model $\mathbf{m}$ are straight lines from the seismic source in (X_s, Y_s, Z_s) to the known position of each seismic sensor. Thus, predicted arrival times $t^{\text{pred}}_{x,i}$ (with $x = P, S$ and $i = 1, ..., N$) can be easily computed from the source-sensor distance, V_P, and V_P/V_S. The vector $\mathbf{e} = t^{\text{obs}}_{x,i} - t^{\text{pred}}_{x,i}$ contains the residual differences between predicted and observed arrival times. The hyperparameters π_P and π_S multiply the error variances that define the error covariance matrix as follows: $\sigma_x^2(\mathbf{m}) = \sigma_0^2 \cdot 10^{2\pi_x}$, where $x = P$ or S. Therefore, the error covariance $\mathbf{C}_e(\mathbf{m})$ is a diagonal matrix that contains the values of $\sigma_x^2(\mathbf{m})$ for each P- or S-wave arrival time.

In this context, the likelihood function can be written as

$$L(\mathbf{m}) = P(\mathbf{d} \mid \mathbf{m}) = \frac{1}{[(2\pi)^N |\mathbf{C}_e(\mathbf{m})|]^{1/2}} \exp\left(-\frac{1}{2}\mathbf{e}^T \mathbf{C}_\mathbf{e}(\mathbf{m})^{-1}\mathbf{e}\right), \tag{1}$$

Model parameter	Name	Minimum	Maximum	Scale
X-coordinate of source	X_s	0 m	5000 m	0.05
Y coordinate of source	Y_s	5000 m	6900 m	0.05
Z coordinate of source	Z_s	-1600 m	-200 m	0.15
Origin time of source	OT_s	-0.2 s	0.2 s	0.05
P-wave velocity	V_P	4.0 km/s	8.0 km/s	0.10
V_P/V_S ratio	V_P/V_S	1.5	1.9	0.20
Hyperparameter (P-wave)	π_P	-0.5	5.0	0.075
Hyperparameter (S-wave)	π_S	-0.5	5.0	0.075

Table 1. Prior probability distributions of model parameters in vector $\mathbf{m}$. The "Scale" column lists the multipliers of the prior ranges that define the standard deviation of the normal probability distributions used to generate candidate models in the Metropolis algorithm.

Model parameter	Name	Minimum	Maximum	Scale
Number of shells	k	1	100	–
Shell radii	$\mathbf{r}_k$	0 m	4000 m	0.02
Shell weights (P-wave)	$\mathbf{w}_{k,P}$	0.0	3.0	0.02
Shell weights (S-wave)	$\mathbf{w}_{k,S}$	0.0	3.0	0.02

Table 2. Prior probability distributions of spherical shell parameters in vector $\mathbf{m}$. The "Scale" column lists the multipliers of the prior ranges that define the standard deviation of the normal probability distributions used to generate candidate models in the Metropolis algorithm.

where the determinant of the diagonal covariance matrix is the product of its diagonal entries, i.e., $|\mathbf{C}_e(\mathbf{m})| = \sigma_P^{2N_P}(\mathbf{m}) \cdot \sigma_S^{2N_S}(\mathbf{m})$.

For each simulation, we ran 10 independent sampling chains starting from 10 different model vectors randomly sampled from the prior distribution. Each chain ran for 10^6 sampling iterations, with a $5{\cdot}10^5$ initial burn-in period. We save 1 every 1000 models, so that our final posterior distribution is reconstructed from $5{\cdot}10^4$ model vectors. Each simulation took less than 60 seconds on a laptop (Apple M1 CPU), but Message Passing Interface (MPI) directives can reduce the CPU time to few seconds.

1.4 Assigning data weights with a trans-dimensional algorithm.

The novel approach applied in our second and third experiments (using arrival times from the calibration blast and the natural event) follows the algorithm presented in (Piana Agostinetti & Sgattoni, 2021), except that here we use spherical shells in space rather than change points in time. In our case, the spherical shells are centered at an approximate preliminary location of the source (see below for a discussion on determining a preliminary source location). The models sampled by McMC are composed by six physical parameters, as done earlier, plus a variable number of parameters related to the spherical shells: the number of shells k, the k-vector of shell radii $\mathbf{r}_k$, and the k+1-vectors of the weights for P- and S-wave arrivals in each shell $\mathbf{w}_{k,P}$ and $\mathbf{w}_{k,S}$. Thus, $\mathbf{m} = (X_s, Y_s, Z_s, OT_s, V_P, V_P/V_S, k, \mathbf{r}_k, \mathbf{w}_{k,P}, \mathbf{w}_{k,S})$. Priors for the spherical shell parameters are uniform (see Table 2).

Given the model vector $\mathbf{m}$, we compute a modified version of the covariance matrix of the data errors $\mathbf{C_e}(\mathbf{m})$ that accounts for the data weights as follows:

$$\mathbf{C_e}(\mathbf{m}) = \mathbf{W}^{-1}(\mathbf{m})\, \mathbf{C_e^*}\, \mathbf{W}^{-1}(\mathbf{m}) \quad (2)$$

Seismic source	X (m)	Y (m)	Z (m)
Calibration blast	3230.1	6311.7	-1077.4
Natural event	2144.6	6325.9	-1145.6

Table 3. Preliminary event locations used to compute the source-sensor distances δ_i.

where $\mathbf{W}(\mathbf{m})$ is a diagonal matrix that contains the weight assigned to each arrival time; as $\mathbf{C}^*_\mathbf{e}$ is diagonal, $\mathbf{C}_\mathbf{e}(\mathbf{m})$ is also diagonal. We note here that an equivalent view is that the algorithm samples variances, rather than weights, in each of the spherical shells as in a hierarchical Bayes strategy (Malinverno & Briggs, 2004). In fact, in a Bayesian framework, "data noise" includes observational and theoretical errors. Here, as the source-station increases, the theoretical "noise" due to a wrong velocity model increases. The resulting variances in the diagonal matrix $\mathbf{C}_\mathbf{e}(\mathbf{m})$ will be consistent with the size of the residual differences between predicted and observed arrival times in the vector $\mathbf{e} = t^{\text{obs}}_{x,i} - t^{\text{pred}}_{x,i}$.

Writing δ_i as the distance from the source of the sensor recording the i-th arrival time, the entries of $\mathbf{W}$ are

$$\mathbf{W}_{ii} = 10^{-w_i(\mathbf{m})}, \tag{3}$$

where $w_i(\mathbf{m})$ is

$$w_i(\mathbf{m}) = \begin{cases} w_{1,x} & \text{if} \quad \delta_i < \tilde{r}_1 \\ w_{j,x} & \text{if} \quad \tilde{r}_j < \delta_i < \tilde{r}_{j+1} \qquad 1 < j < k \\ w_{k,x} & \text{if} \quad \tilde{r}_k < \delta_i, \end{cases} , \tag{4}$$

where $x = P$ or S depending on whether the i-th arrival time is for a P- or S-wave and $\tilde{\mathbf{r}}_k$ is the ordered version of the vector $\mathbf{r}_k$ (i.e. $\tilde{r}_1 < ... < \tilde{r}_k$). The likelihood function is as in Equation 1 with the covariance matrix $\mathbf{C}_\mathbf{e}(\mathbf{m})$ of Equation 2.

Composing the covariance matrix $\mathbf{C}_\mathbf{e}(\mathbf{m})$ needs a preliminary source location to compute the source-sensor distances δ_i. In general, an approximate preliminary location is available soon after the event takes place and can be safely used. If a preliminary event location was not available, the position of the sensor that receives the earliest P-wave arrival can be used as well. Here we use as preliminary locations the actual location of the calibration blast and the location of the natural event published in (Dineva et al., 2022); the coordinates of these preliminary locations are in Table 3.

Finally, we briefly illustrate the procedure to sample the parameter space. Physical parameters are explored following the same approach as in the standard McMC sampling presented by (Riva & Piana Agostinetti, 2023), whereas sampling of the parameters related to the spherical shells follows the recipe described in (Piana Agostinetti & Sgattoni, 2021). During McMC sampling, candidate models with perturbed physical parameters are generated in the odd-numbered iterations while the spherical shell parameters are perturbed in even-numbered iterations. Given the high number of parameters involved, we ran 20 independent chains of 10^6 iterations each. The full CPU time for the 20 chains is about 200 seconds, but MPI directives can be used to make the chains running in parallel.

References

Dineva, S., Dahner, C., Malovichko, D., Lund, B., Gospodinov, D., Piana Agostinetti, N., & Rudzinski, L. (2022). Analysis of the magnitude 4.2 seismic event on 18 may 2020 in the Kiirunavaara mine, Sweden. *RaSiM Conference,*.

Malinverno, A., & Briggs, V. A. (2004). Expanded uncertainty quantification in inverse problems: Hierarchical Bayes and empirical Bayes. *Geophysics*, *69*(4), 1005-1016. (doi:10.1190/1.1778243)

Piana Agostinetti, N., & Sgattoni, G. (2021). Changepoint detection in seismic double-difference data: application of a trans-dimensional algorithm to data-space exploration. *Solid Earth*, *12*, 2717–2733. doi: 10.5194/se-2021-79

Riva, F., & Piana Agostinetti, N. (2023). *The micro-seismicity of Co. Donegal (Ireland): defining baseline seismicity in a region of slow lithospheric deformation.* (submitted to *Terra Nova*)

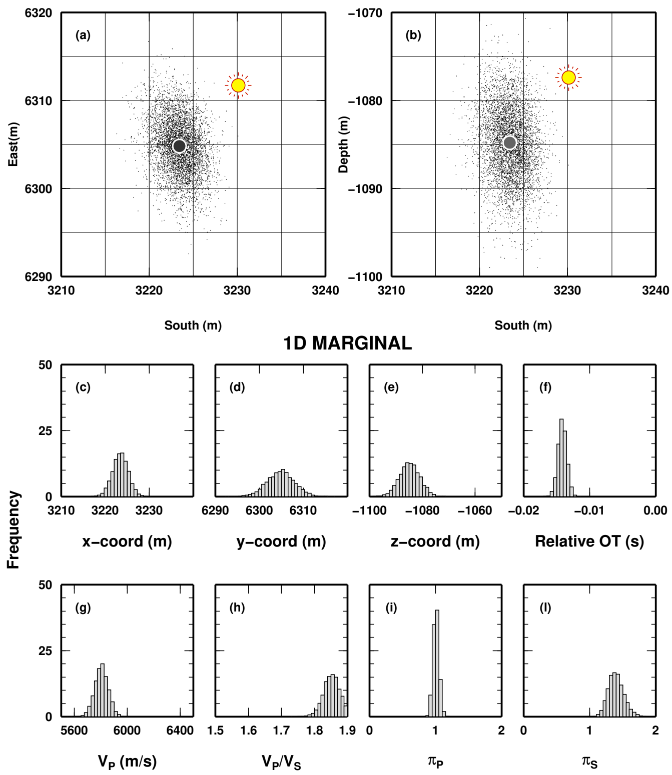

Figure 1. Standard McMC sampling applied to locating the calibration blast in the reference solution when using all arrival times recorded up to 800 m from the blast. (a) Sampled solutions projected onto the horizontal X-Y plane (black dots) and posterior mean (black circle). The yellow sun indicates the true position of the calibration blast. (b) Sampled solutions projected onto the vertical X-Z plane. Symbols as in panel (a). (c-l) Histograms that approximate the posterior distribution of the sampled parameters. "Relative OT" is the origin time of the blast relative to the earliest recorded P-wave arrival time.

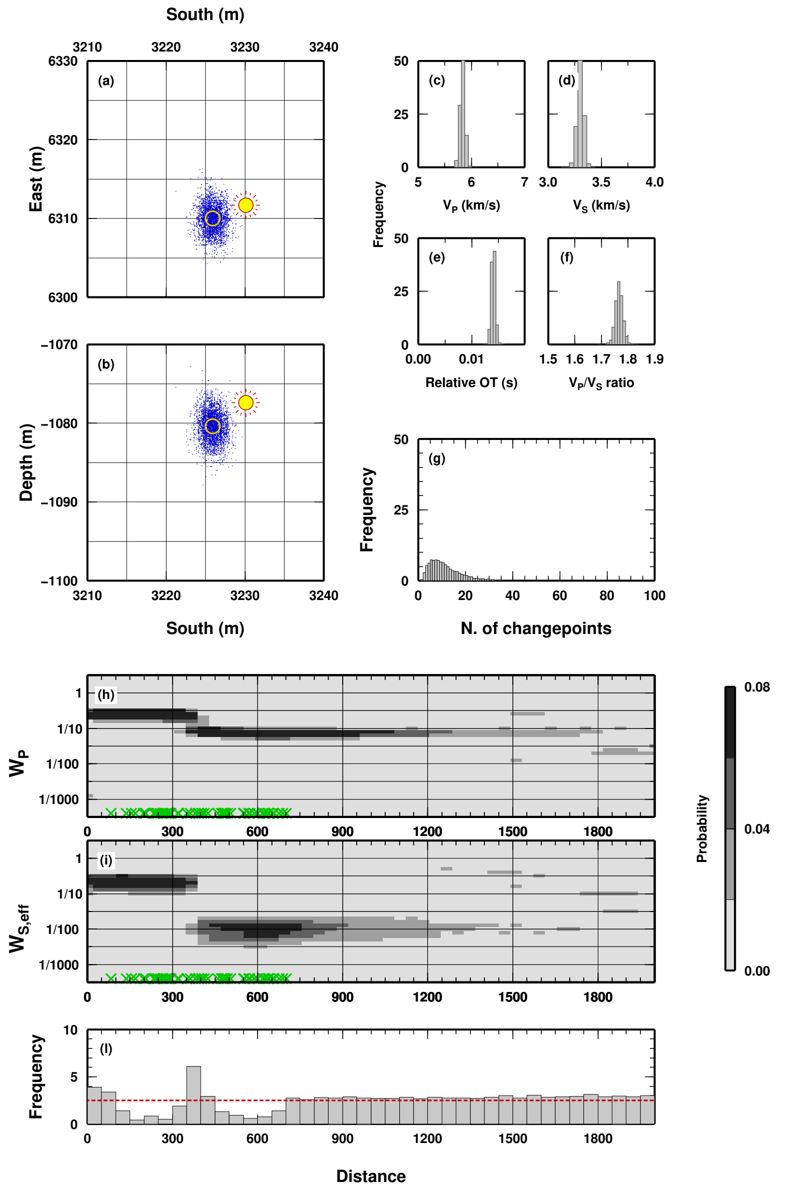

Figure 2. Novel data-weighing method applied to locating the calibration blast. (a) Sampled solutions projected onto the horizontal X-Y plane (blue dots) and posterior mean (blue circle). The yellow sun indicates the true position of the calibration blast. (b) Sampled solutions projected onto the vertical X-Z plane. Symbols as in panel (a). (c-g) Histograms that approximate the posterior probability density function (PDF) of the sampled parameters. "Relative OT" is the origin time of the blast relative to the earliest recorded P-wave arrival time. (h) Posterior PDF of the weights assigned to P-wave arrival times as a function of source-sensor distance. Green crosses indicate the distance of each sensor from the source. (i) As in (h) for S-wave arrival times.(l) Posterior PDF of shell radii. The red dashed line indicates the uniform prior pdf.

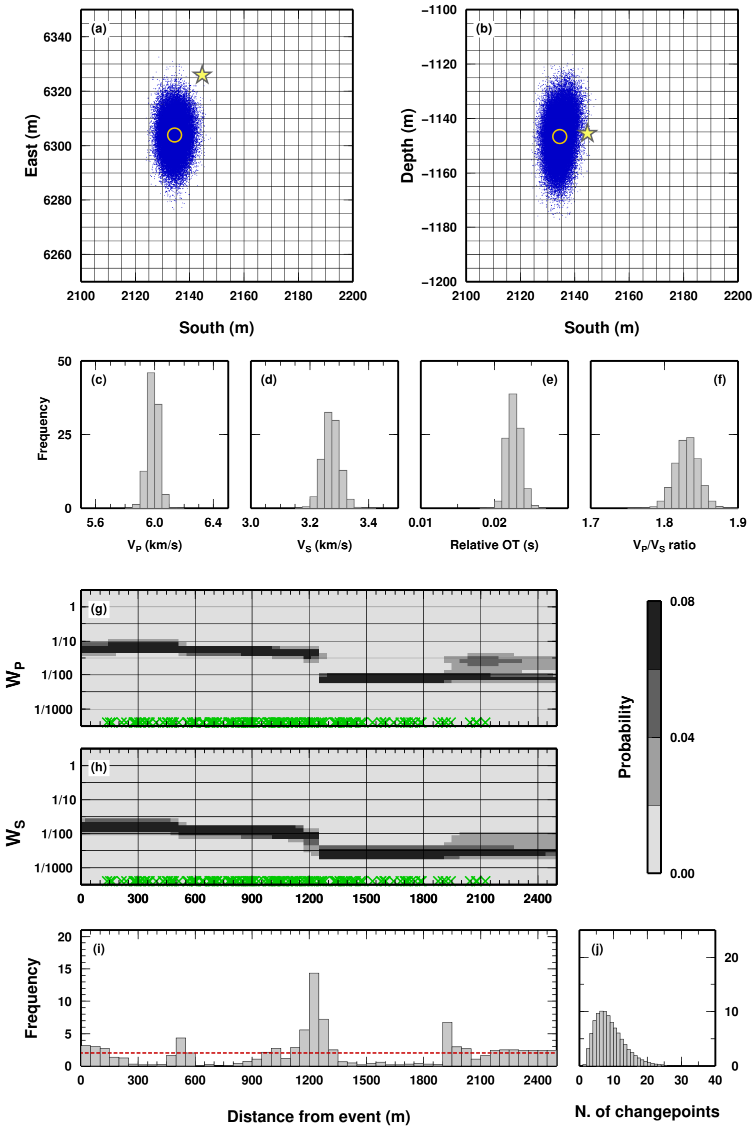

Figure 3. Application of the novel data-weighing method to recordings of the M_w 4.2 natural event. See the caption of Figure 2 for a description of panels and symbols. The yellow star indicates the preliminary position of the natural event.

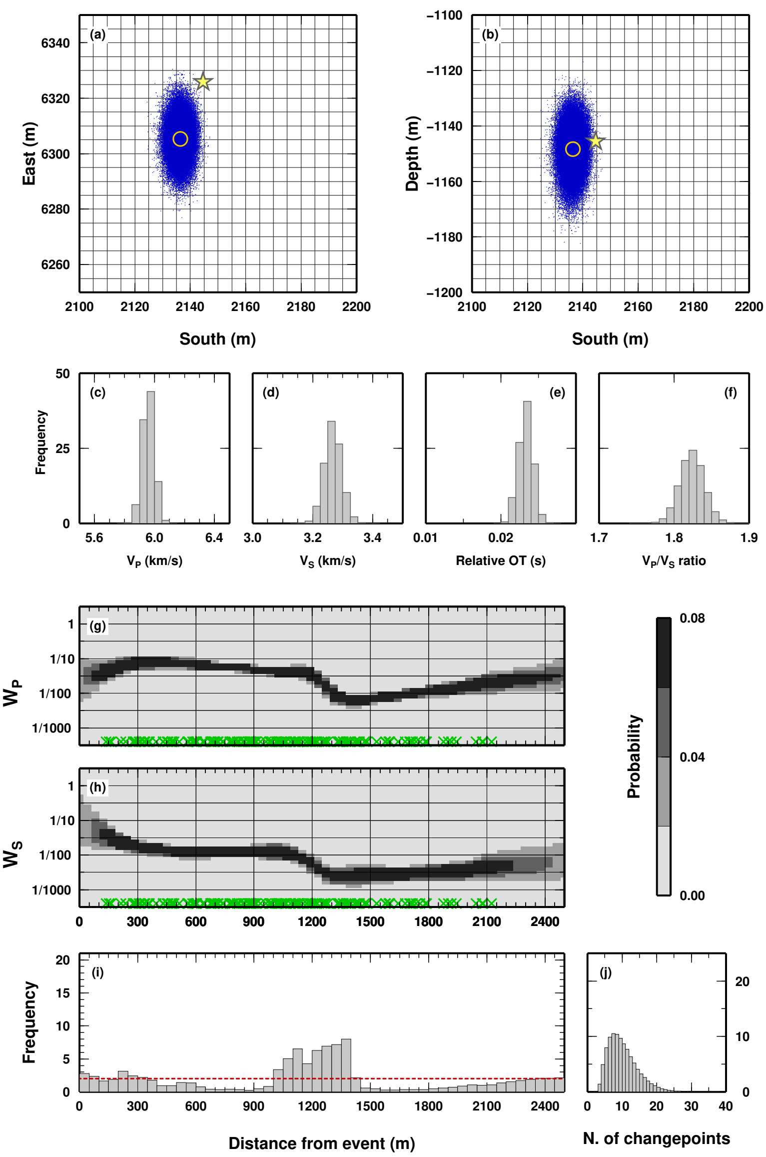

Figure 4. Application of the novel data-weighing method to recordings of the M_w 4.2 natural event when using the alternative parameterization with linearly varying weights within each spherical shell (Figure **??**c). See the caption of Figure 2 for a description of panels and symbols. The yellow star indicates the preliminary position of the natural event.