Tectonics

Supporting Information for

The M7 Samos 2020 earthquake: a snapshot of transtension within the Samos - Ikaria Basins

V. Plicka*[1], F. Gallovič[1], J. Zahradník[1], A. Serpetsidaki[3], E. Sokos[3], N. Vavlas[2], and A. Kiratzi[2]

[1] Charles University, Faculty of Mathematics and Physics, Czech Republic

[2] Aristotle University of Thessaloniki, Department of Geophysics, Greece

[3] University of Patras, Department of Geology, Seismological Laboratory, Greece

*Corresponding author: Vladimír Plicka (vp@karel.troja.mff.cuni.cz)

Contents of this file

This supplement includes supporting information on the data and methods used together with additional results. Each section consists of the text and the corresponding figures.

Section 1 - Hypocenter relocations

A. Mainshock relocation

We relocated the hypocenter of the mainshock, using phase arrivals from broadband and strong motion stations from regional networks (see Data and Resources). We handpicked 81 P- and 40 S-wave arrival times at 81 stations at epicentral distances from 20 to 290 km (Fig. S1). Using a subset up to 160 km (circle in Fig. S1), we obtained similar results (epicenter differs ~+-1 km from the location from all stations). We excluded any S-phase arrivals with large residuals or unclear arrivals; S-phases from the nearest stations are included to constrain the hypocenter depth. We selected eight velocity models (VM) (Akyol et al., 2006, Crust1.0 – Laske et al., 2013, Kalafat, 1987, Kayapack and Gokkaya, 2012, Konstantinou, 2018, Novotný et al., 2001, Özer and Polat, 2017, Özer et al., 2018) as most suitable for this application (Fig. 2a, b in the main text). Based on the preliminary analysis, we excluded the Kaypak and Gökkaya, 2012 and Kalafat, 1987 velocity models because they provided inconsistent results compared to the others (very deep hypocentres, ~30km and ~18km for Kaypak and Gökkaya, 2012 and Kalafat, 1987, respectively). For the models that do not provide Vs velocities, i.e., Akyol et al. (2006) and Özer and Polat (2017), we used Vp/Vs=1.75 obtained from the mainshock dataset and Vp/Vs=1.73 for the Konstantinou (2018) model as proposed in his work.

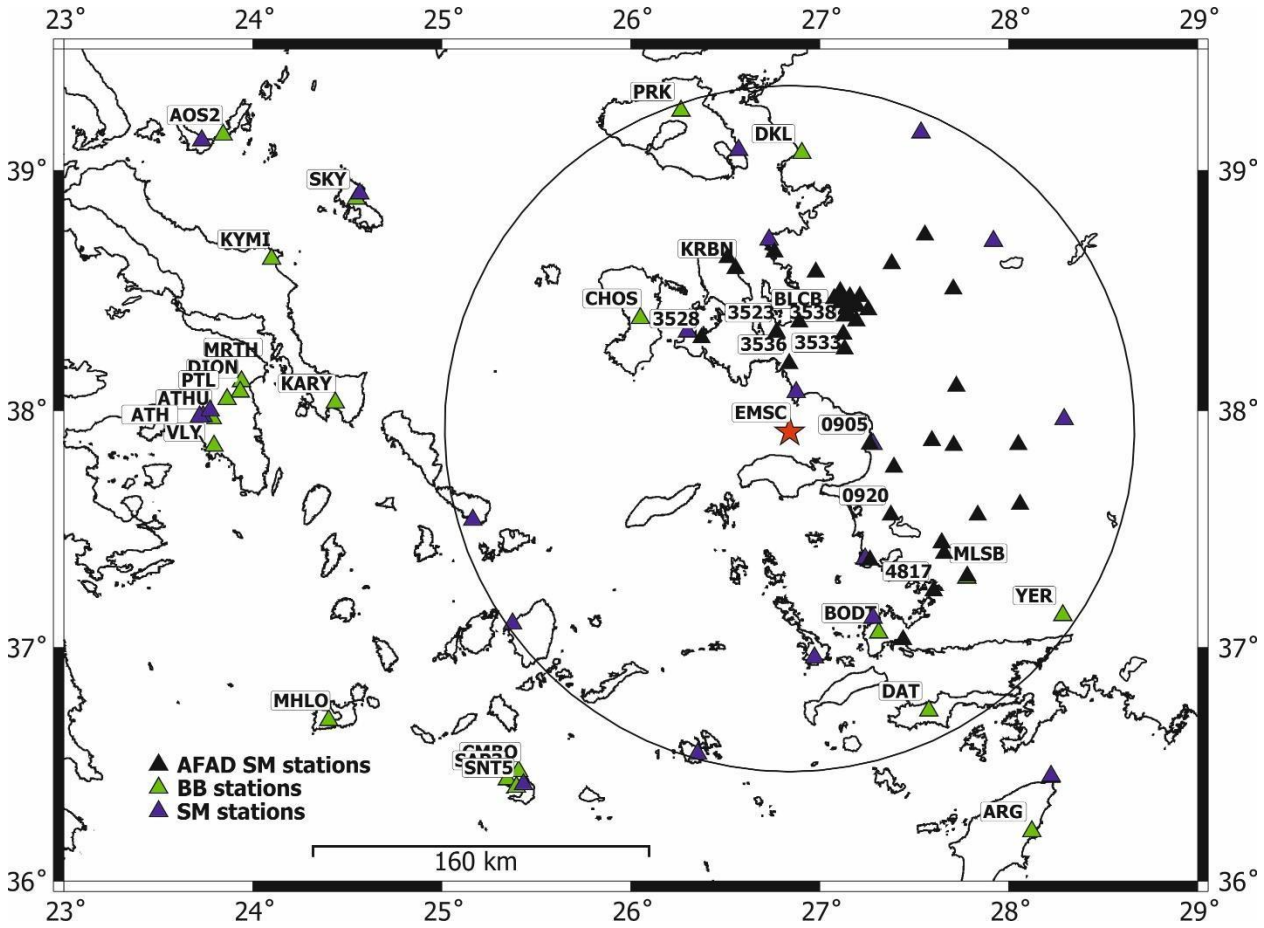


Figure S1. Location (triangles) of broadband and strong motion stations used in NonLinLoc relocation of the mainshock (star). The stations within the circle of radius of 160 km were used as an alternative subset. Using this subset does not significantly change the location of the mainshock. The locations in the Table 1 are obtained from all stations.

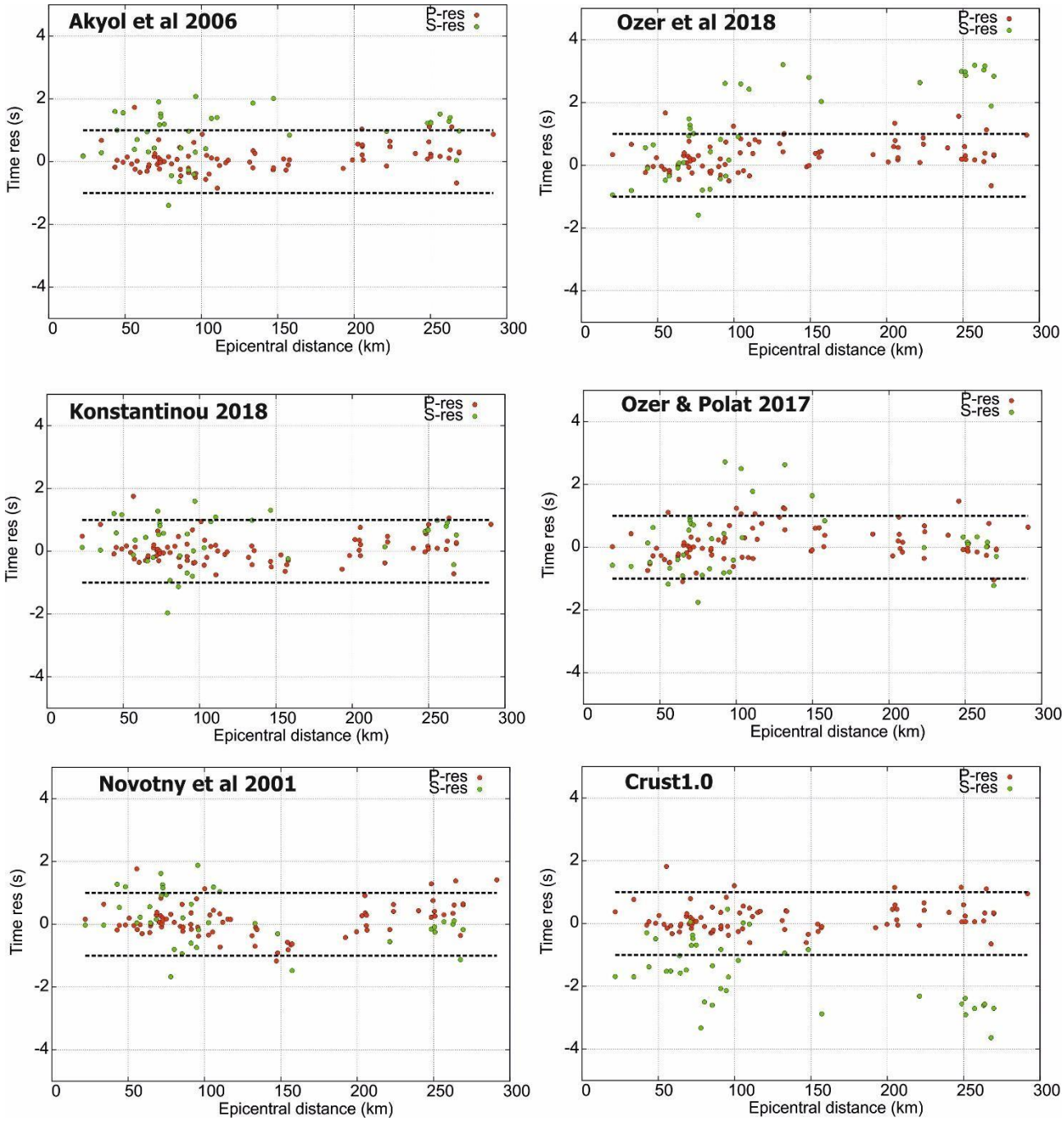


Figure S2 Residuals for P and S-wave arrivals as a function of epicentral distance for the different models tested to relocate the mainshock.

B. Relocation of the aftershocks

Manual arrival time picks of events recorded in the stations of Fig. S3 during the first 40 days were used in the initial location

The Vp/Vs ratio was set equal to 1.69 based on the Wadati diagram (Fig. S4). Results for eight models are shown in Fig. S5. The comparison was initially made using the HYPOINVERSE location errors and the hypocenters' distribution. The crustal model of Özer et al., 2018 was finally selected since it depicted the lowest data misfits (i.e., mean RMS ~ 0.31s) and has been derived from a most recent seismic experiment in the study area.

Subsequently, the double-difference relative relocation HYPODD (Waldhauser and Ellsworth 2000; Waldhauser 2001) procedure was used combining the P- and S- wave arrival times (84500 phases) derived from stations within 100 km from the mainshock's epicentral area (Fig. S6). HYPODD approach improves the location accuracy by reducing the influence of the inaccuracy of existing velocity models. The double-difference residuals for the pairs of earthquakes at each station were minimized by weighted least squares, using the method of conjugate gradient least squares (LSQR). Errors reported by LSQR are grossly underestimated and need to be assessed independently by using the singular value decomposition technique (SVD) on a subset of events (Waldhauser& Ellsworth 2000). Therefore, we relocated a subset of 100 events from the first day of the sequence using the SVD method. A good data fit, mean RMS < 0.1 s, and relatively low errors were obtained for longitude/latitude < 0.5 km, and depth < 1 km). The velocity model used in the relocation was the model used in the initial location process. More than 1300 events were relocated and clustered in the area of interest. The HYPODD final results include 79.5% of the initial dataset, showing a spatial pattern more compact compared to

HYPOINVERSE. The relocated events are more densely concentrated than the initially located ones in one major and three minor clusters (Fig. S6).

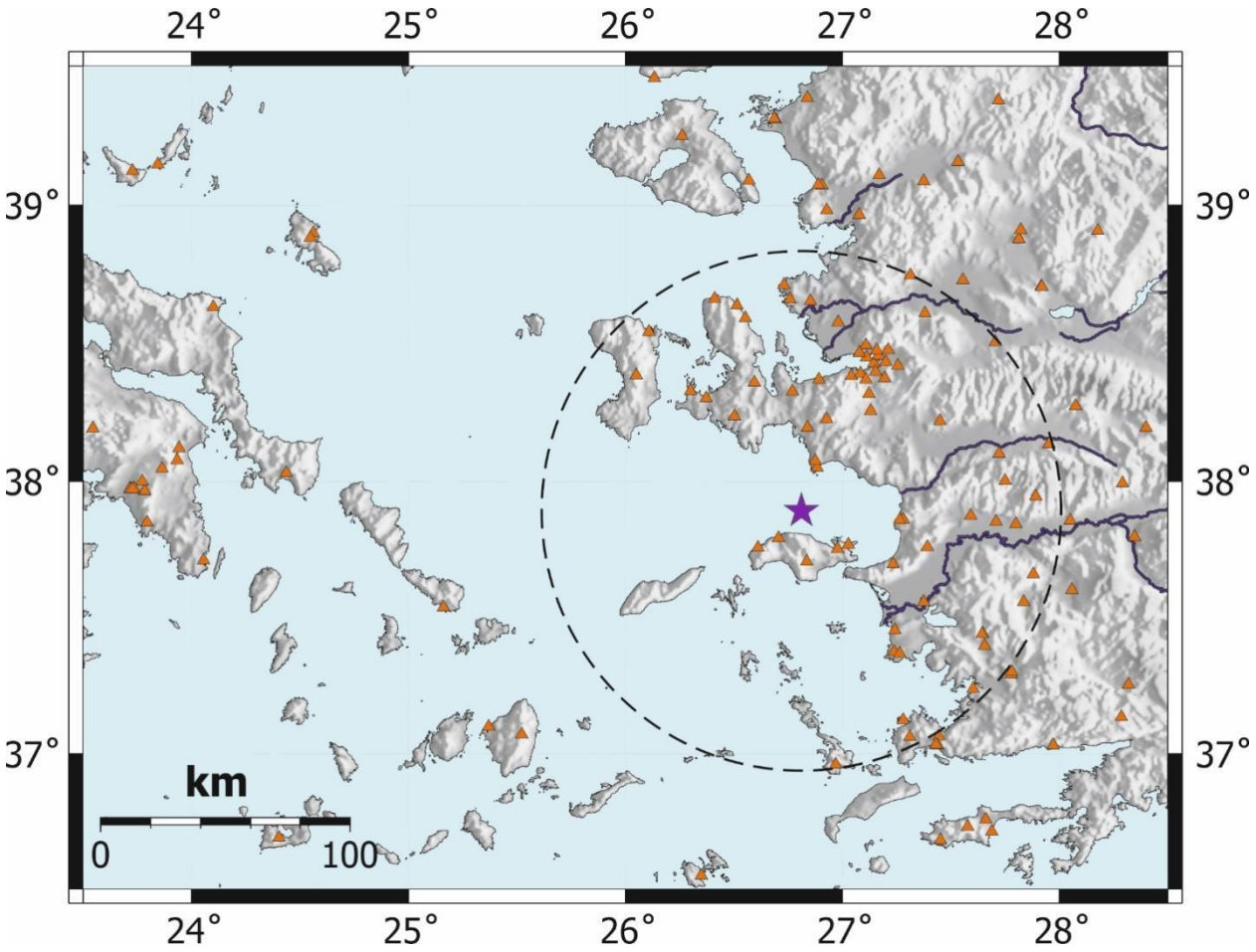


Figure S3. Distribution of stations (triangles) used for the relocation of the aftershocks. The dashed circle encloses the stations located within 100 km from the epicenter (star).

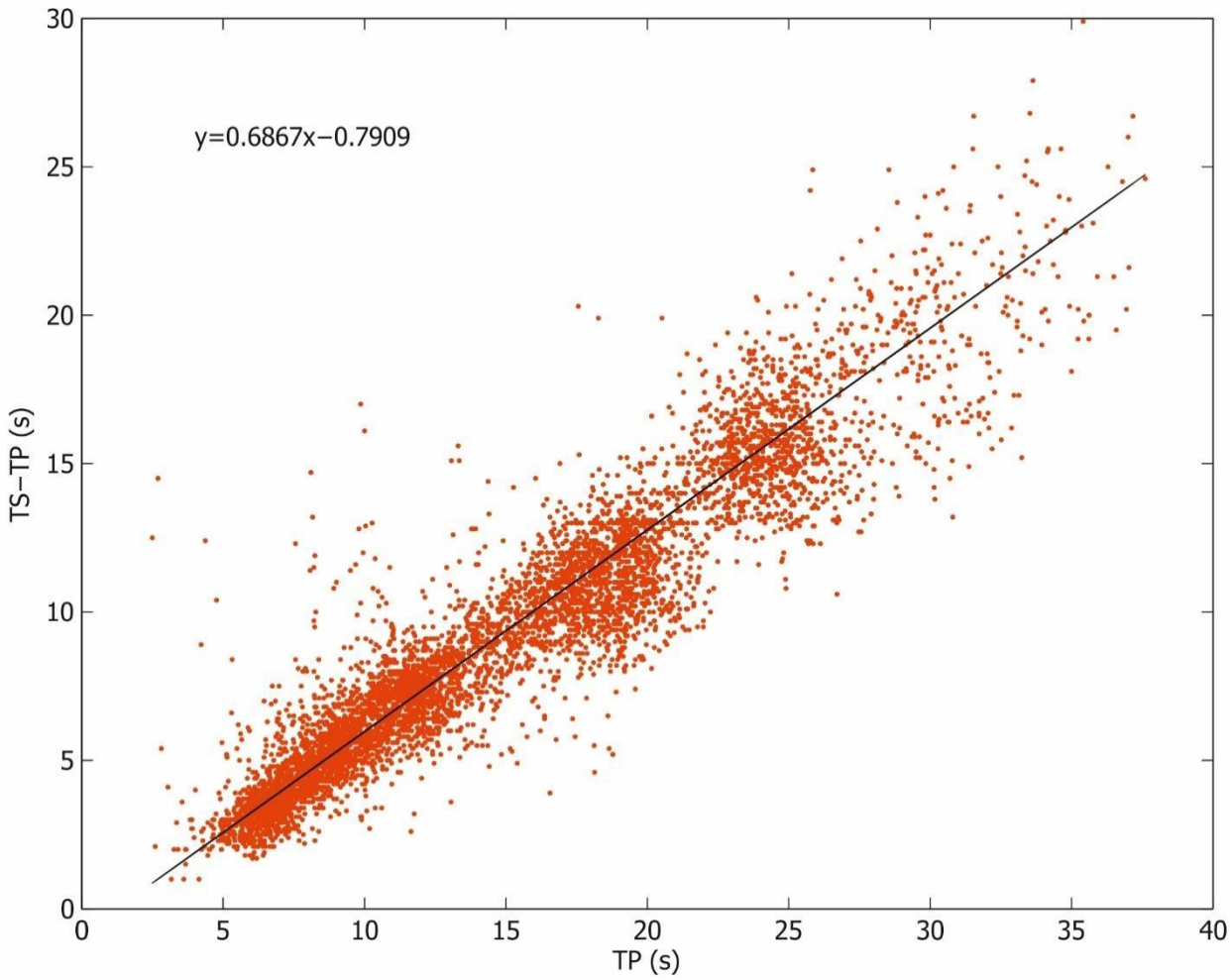
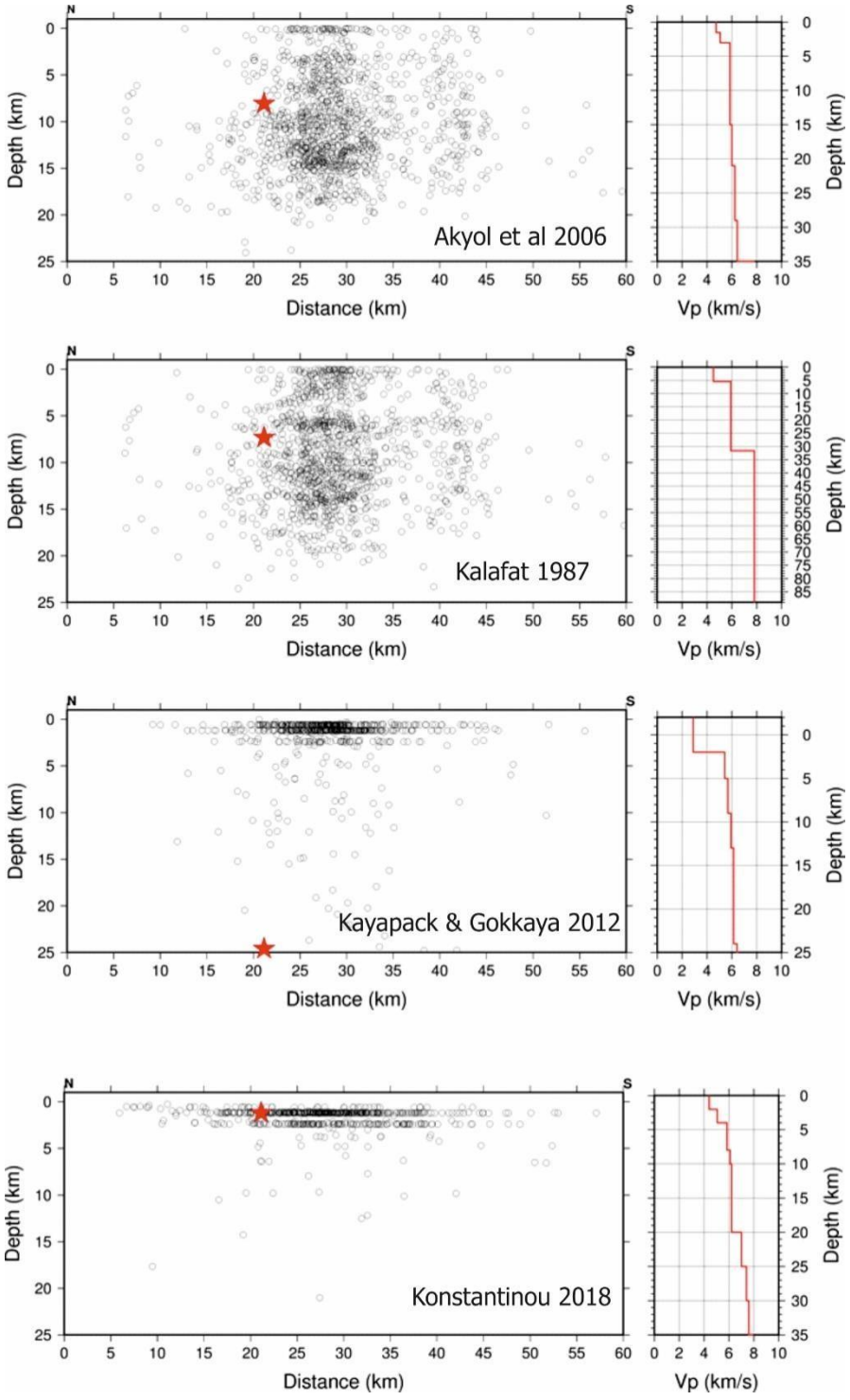


Figure S4. Wadati diagram to calculate the Vp/Vs ratio of the aftershock sequence


N
S
Depth (km)
Distance (km)
Vp (km/s)
Akyol et al 2006
Kalafat 1987
Kayapack & Gokkaya 2012
Konstantinou 2018

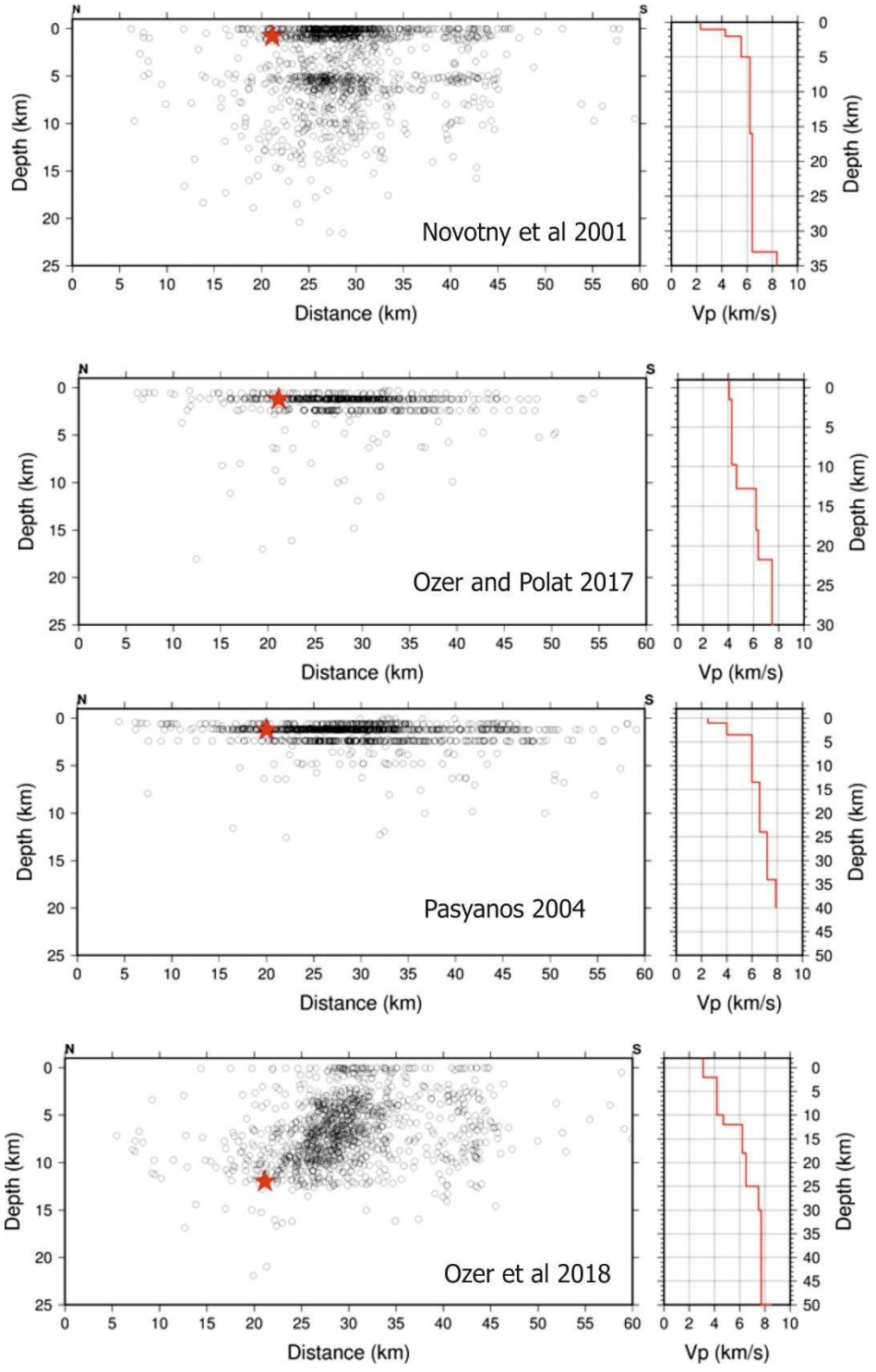


Figure S5. Relocation of the sequence in different velocity models applicable to the region. The preferred model is the Özer et al. (2018), finally adopted for the relocation.

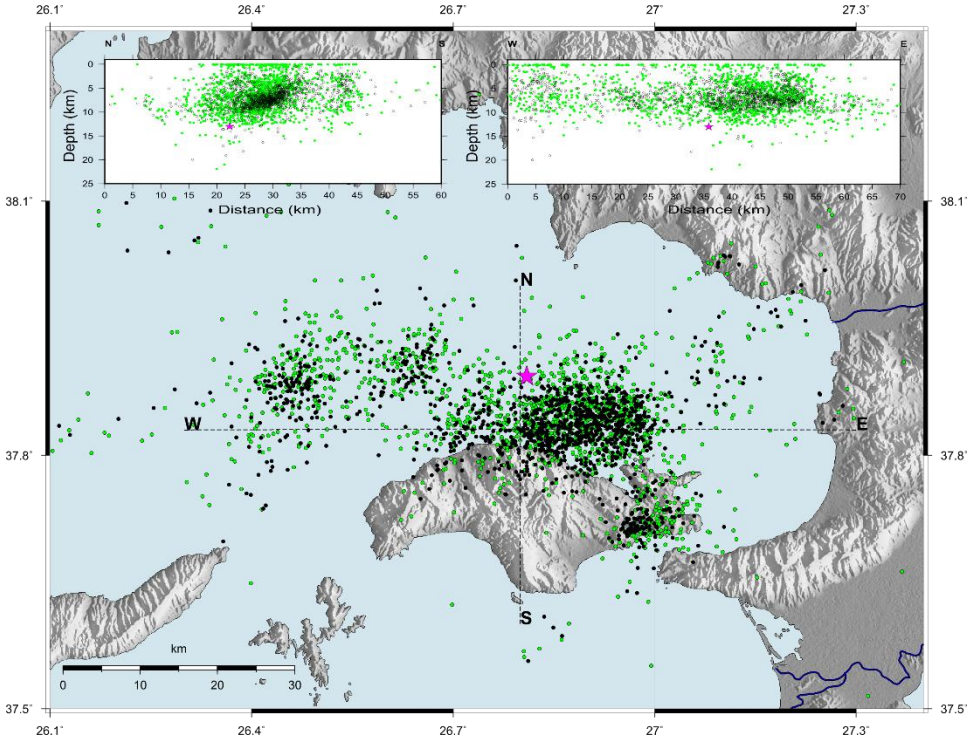
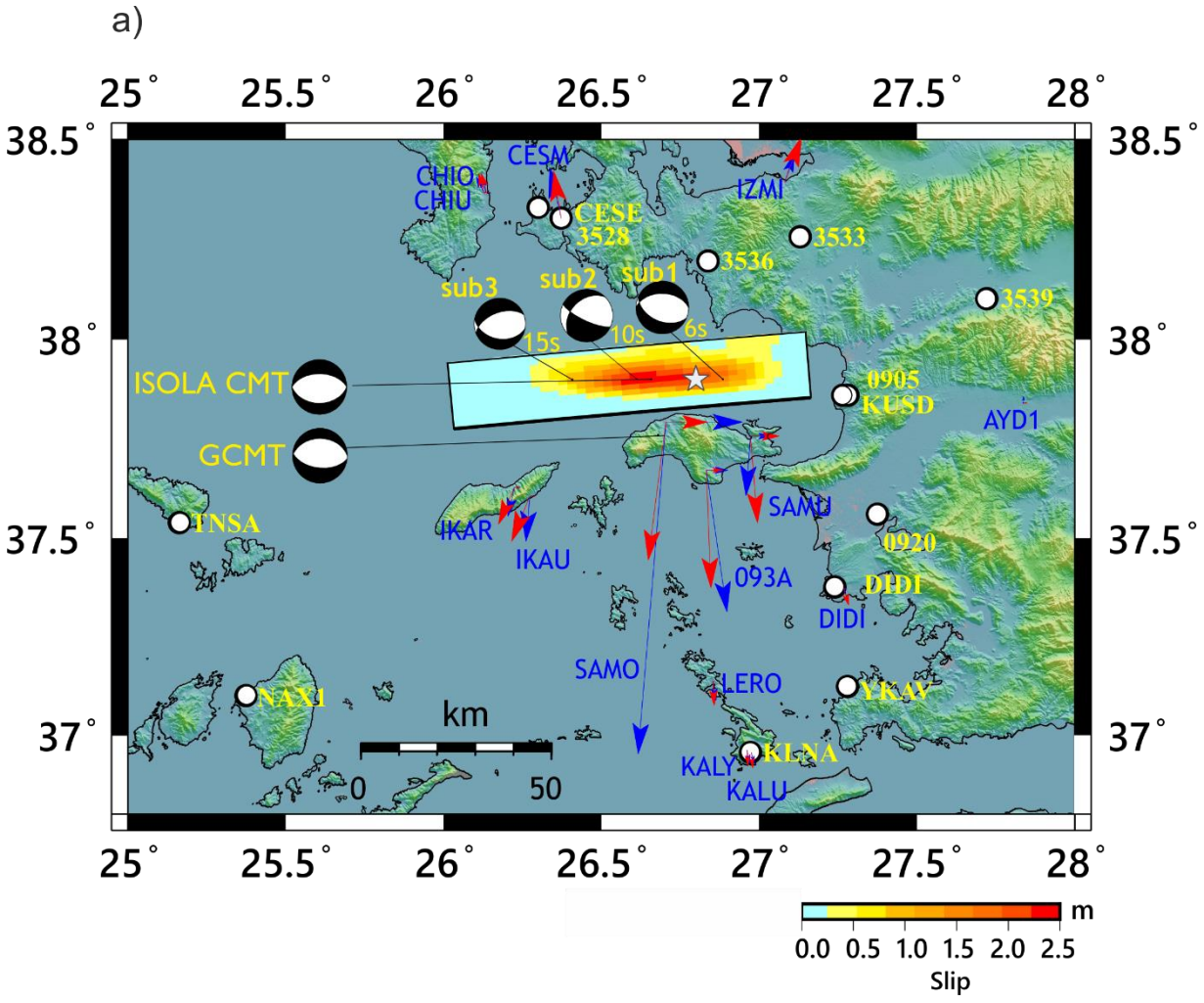


Figure S6. Comparison of initial (green dots) and relocated aftershock hypocenters (black dots) alongside the designated cross-sections. The data cover the period October 30 to December 4, 2020. The star symbol represents the relocated mainshock.

Section 2. Finite–Fault kinematic rupture model

Here, we show the results of the LinSlipInv kinematic slip inversion without including the GNSS geodetic data. Note the deficiency of the slip model in the shallow slip patch that reached the surface and the misfit on the predicted synthetic displacement on the Samos inland station SAMO, located in the area where the maximum uplift was measured and where the tsunami initiated. The localized shallow slip in our preferred model (Fig. 4) is crucial in fitting the observed 35.7 cm static displacement in SAMO.



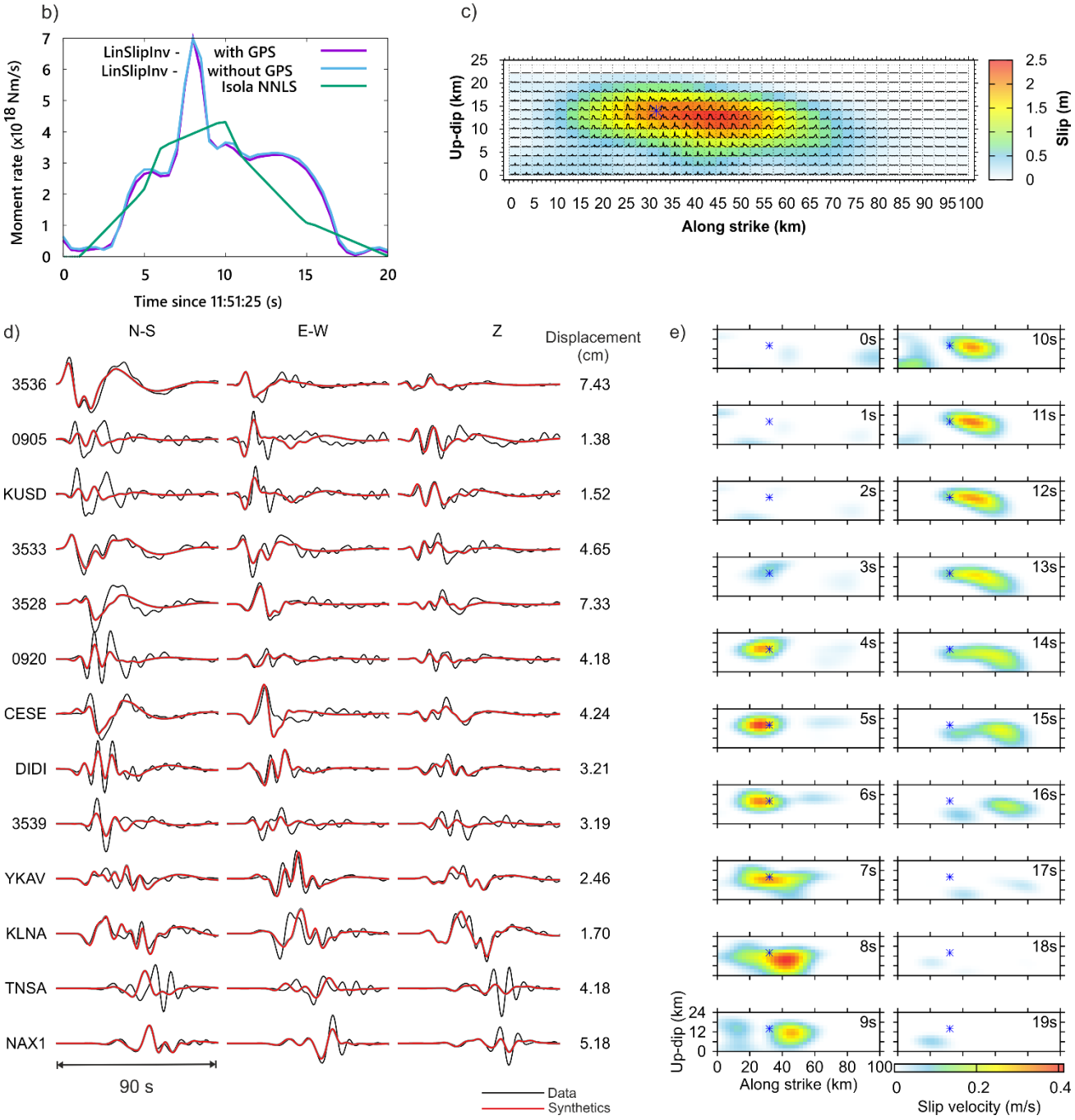


Figure S7. a) Map view of slip distribution when the geodetic displacements are not included in the inversion. Note the misfit between the GNSS daily solution (blue) and the predicted synthetic displacements (red) at SAMO. b) Moment rate functions. c) Slip distribution (with slip rate functions superimposed (maximum scaled to 0.5m/s). d) Waveform fit of displacement recordings. Synthetics (grey) from the inversion of the main text (when the geodetic data were included) almost perfectly coincide with the synthetics (red) without geodetic data, suggesting no resolution power of the seismic data alone regarding the

shallow slip. d) Slip rate snapshots of the model (Fig. S7a,c). All other notations as in Figures 4 and 5 of the main text.

Section 3. New Empirical Green's function (EGF) method to calculate Apparent Source Time Functions (ASTFs)

The waveforms *s(t)* and *S(t)* of the weak event (EGF) and the mainshock, respectively, are defined by Eqs. (1) and (2):

$$s(t) = m(t) * g(t) \quad (1)$$

$$S(t) = M(t) * g(t) \quad (2)$$

Green's function $g(t)$ is the same for both events and need not be known. The $m(t)$ and $M(t)$ are the moment rate functions. We assume a frequency range (detailed below) in which $m(t)$ can be approximated as an isosceles triangle, centered at time $t = 0$, whose duration is shorter than the duration of $M(t)$. Function $M(t)$ is expressed as a set of equidistantly shifted functions $m(t)$; see Eq. (3), where w_i are the unknown weights. The time shift values $\tau_i = (i-1)\Delta\tau$ and their number N are predefined.

$$M(t) = \sum_{i=1}^{N} m(t-\tau_i) w_i \quad (3)$$

Thus the mainshock $S(t)$ can be represented as a weighted sum of the shifted EGF records:

$$S(t) = \left[\sum_{i=1}^{N} m(t-\tau_i) w_i\right] * g(t) = \sum_{i=1}^{N} s(t-\tau_i) w_i \quad (4)$$

The ratio of the scalar moments of the mainshock and the EGF event (the relative moment) provides a constraint for the weights.

$$M_0 = \sum_{i=1}^{N} m_0 w_i, \ \frac{M_0}{m_0} = \sum_{i=1}^{N} w_i \quad (5)$$

Generalizing for a three-component station (total number of time samples *M*), Eq. (4) with real data *S*, and Eq. (5) yield a system of linear algebraic equations for the weights, see Eq. (6).

$$\begin{pmatrix} s(t_1-\tau_1) & s(t_1-\tau_1) & \cdots & s(t_1-\tau_N) \\ s(t_2-\tau_1) & s(t_2-\tau_2) & \cdots & s(t_2-\tau_N) \\ . & . & . & . \\ s(t_M-\tau_1) & s(t_M-\tau_2) & \cdots & s(t_M-\tau_N) \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ . \\ w_N \end{pmatrix} = \begin{pmatrix} S_1 \\ S_2 \\ . \\ S_M \\ \frac{M_0}{m_0} \end{pmatrix} \qquad (6)$$

In practice, the last row of the matrix and the last value of the data column must be multiplied by a constant, i.e., $const \sum w_i = const \frac{M_0}{m_0}$. The constant does not alter Eq. (5) but guarantees its proper balance with eq. (4). The numerical value of the constant depends on values of S relative to M_0/m_0, depending on whether data S are, e.g., in counts or m/s. In this paper, we use the const. ~ 1.0e7.

Further assuming that $w_i \geq 0$ for each i, we solve the system (6) by nonnegative least-squares inversion (NNLS) after Lawson & Hanson (1974). The inversion quality is measured by the fit between the mainshock recordings and synthetics (Eq. 4), quantified by variance reduction. Eq. (3) then provides $M(t)$ - the desired nonnegative ASTF for a given station.

The procedure is performed on the $s(t)$ and $S(t)$ records equally filtered with a band-pass filter (Harris, 1990). Assuming frequency band (*Fmin, Fmax*), the duration of the triangle $m(t)$ is defined as *1/Fmax*. As such, the shortest temporal variation of the ASTF that can be resolved is *1/Fmax*.

A Fortran code and Gnuplot graphics scripts were developed to perform the inversion and automatically visualize the results. In the case of quality data, the software provides an ASTF, which is:

1. Nonnegative (by definition).
2. Causal, i.e., starting generally at origin time ($t = 0$); for discussion of possible small signals before $t = 0$, see below.

3. Stable, i.e., generally having only minor artifacts beyond the major ASTF part. For details about the artifacts, see below.

4. The area of ASTF is proportional to the relative moment $\left(\frac{M_0}{m_0}\right)$ at each station.

The program reads the three-component ASCII waveforms (time, NS, EW, Z) of the mainshock and EGF, recorded by the same instrument. No instrumental correction is needed. Before the first code run, both seismograms are aligned to have the same P-wave arrival times. Since the mainshock and EGF locations are not identical, the P-wave alignment does not guarantee the S-wave alignment. Therefore, if inverting the whole record, or only S waves, we must allow for a possible start of the resulting ASTF before $t = 0$, mentioned above as the small acausal effect. Regarding compactness of the ASTF, code is executed repeatedly, using either the whole set of the calculated weights $w_{1,\dots N}$, or just $w_{J,\dots K}$, where $J \geq 1$ and $K \leq N$. If the fit between real and synthetic seismograms is similar for the $< J, K >$ interval of the weights, the weights outside of this interval are considered noise (artifact).

The entire inversion process is controlled by a single configuration file. The user can set up several parameters:

a) Using the whole seismogram or defining a (smoothly tapered) time window for inversion that contains, e.g., P or S waves only.

b) Selection of the station components to be used in the inversion.

c) Time interval $T =< t_1, t_2 >$, where weights are to be calculated. It must be greater than the largest expected ASTF duration. A short time interval before origin time (i.e., $t_1 < 0$) is advisable.

d) Time shifts $\Delta\tau$ of the weights w_i. The number of unknowns N in Eq. (6) is given by the time interval T and time shift $\Delta\tau$, $N = {}^{T}/_{\Delta\tau}$.

e) Scalar seismic moments, M_0, m_0.

f) Selecting whether to invert the original or integrated records.

g) Parameters of the band-pass filter.

Graphical outputs:

In this section, the EGF event is Event 1 of the main text; in Fig. S8 we show both Event 1 and 2.

Two Gnuplot scripts (seismo.gpl and rstf.gpl,) display the results. The examples of the output are in Figs. S9 - S12.

The three-component normalized waveforms and amplitude spectra of the mainshock and EGF event are displayed by seismo.gpl Gnuplot script (Figs. S9, S10). The numbers at the waveforms panel are the true amplitudes. The waveforms are filtered in the same frequency range as that used in the inversion. The frequency range is marked by the green zone in the amplitude-spectra plot.

The output of rstf.gpl script (Figs S11, S12) contains the header which summarizes the general parameters (station name, original or integrated input records, frequency band in Hz, and seismic moment ratio) and shows the legend for the waveform panel.

The main result of inversion – i.e. the weights are shown in the top left panel in the time interval $T =< t_1, t_2 >$. The figure below displays the moment rate function constructed following Eq. (3) using the elementary triangles of width =*1/Fmax*, which is shown in the legend together with the maximum moment rate value. The next plot shows the cumulative sum of weights, which is the moment function.

Synthetic seismograms at the right panel are calculated according to Eq. (4) for all estimated weights (red) and their subset $w_{J,\dots K}$ (green) mentioned above. The numbers are the true amplitudes and variance reductions for each component. The user can compare how the subset of the weights fits the observed mainshock (blue) relative to the fit employing all

weights. In this way, redundant weights are identified, and the corresponding 'tail' of ASTF is removed as a noisy artifact.

Amplitude spectra of the NS components for EGF (black) event, observed mainshock (blue), and synthetic mainshock (red) are placed at the left bottom part of the figure.

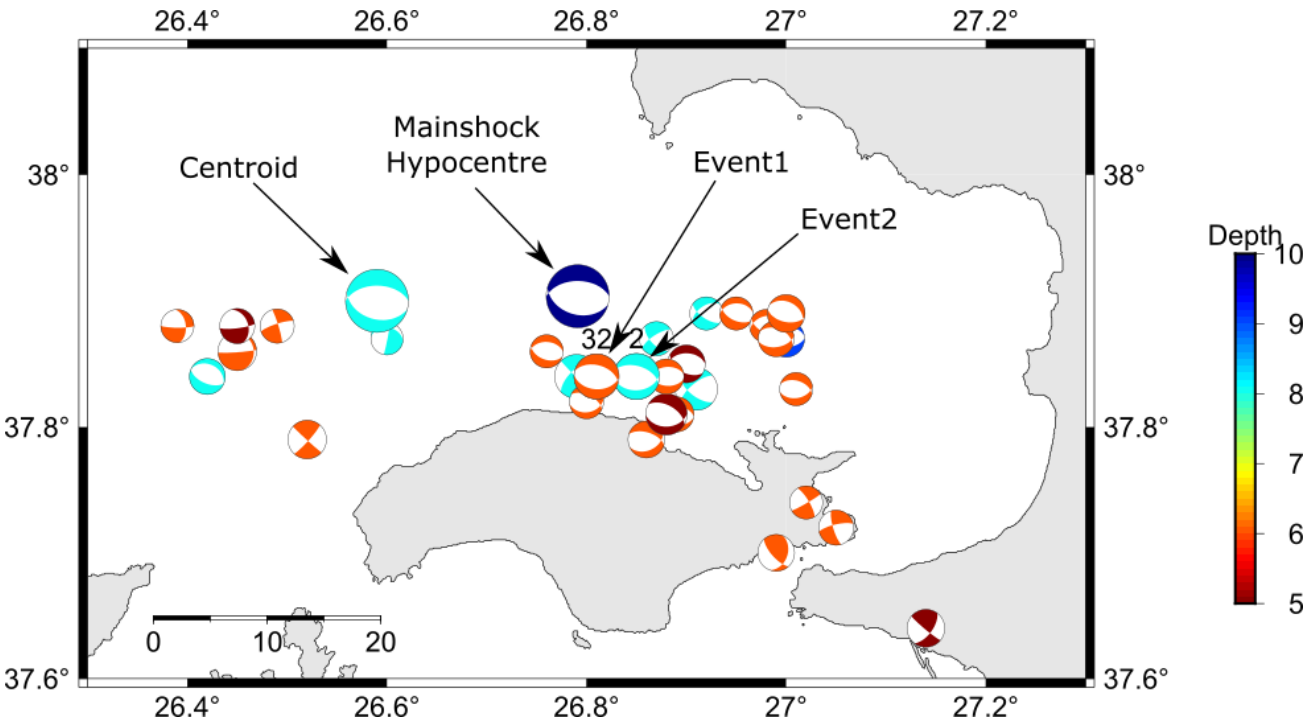


Figure S8. Selection of a suitable EGF: Centroid Moment Tensor (CMT) solutions from Cetin et al. 2020, chapter 1. Only the aftershocks within the depth interval from 5 to 9 km are plotted. The mainshock centroid (see text) and the mainshock NonLinLoc hypocenter (Özer et al., 2018 model) are also depicted. The beach-balls labeled 32 and 2 are selected EGFs (Event1, M5 of October 31, 2020, 05:31 UTC, and Event 2, M5.1 of October 30, 2020, 15:14). Numbers 32 and 2 corespond to the numbers in the Table S3 of Cetin et al., 2020 chapter 1.

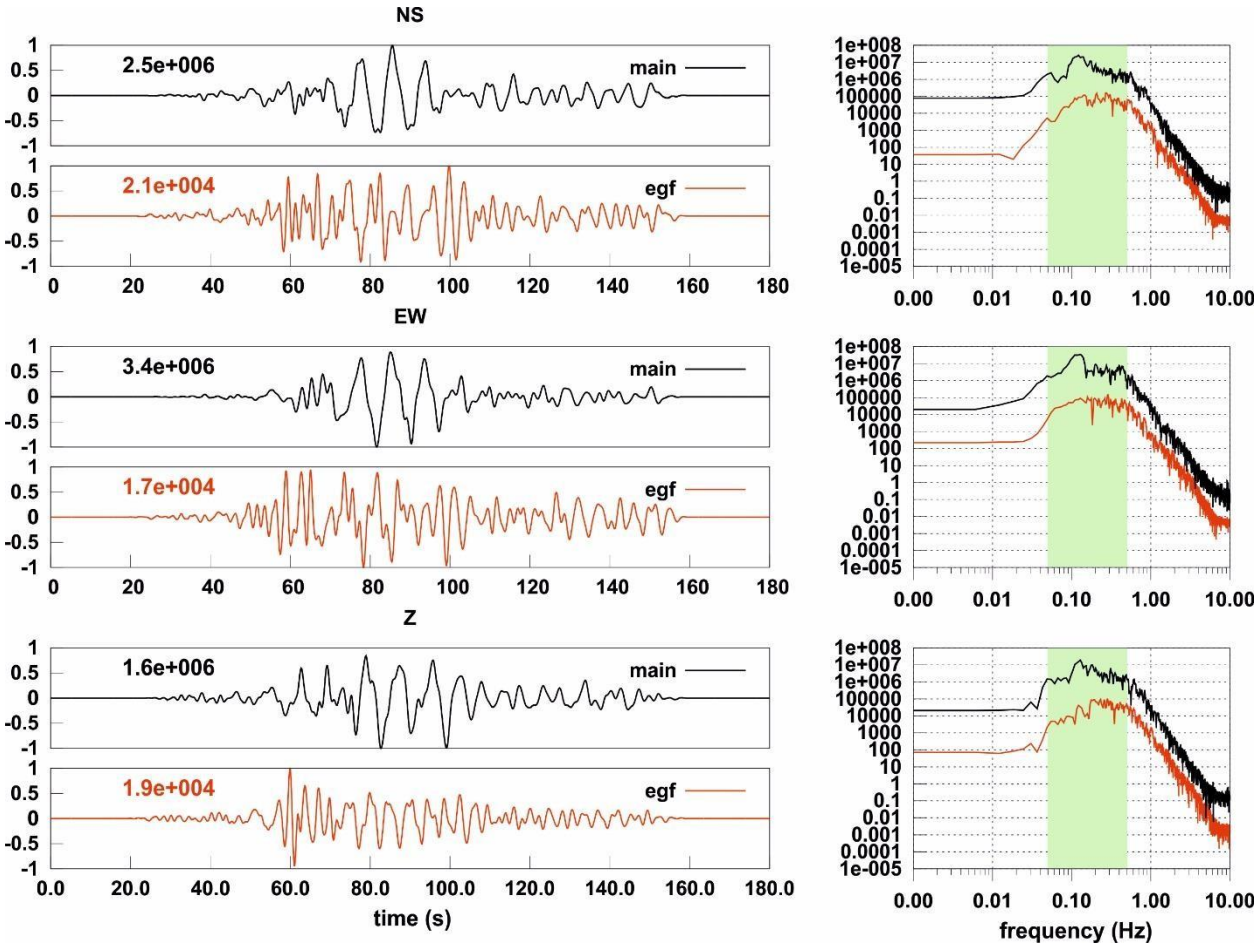


Figure S9. Normalized waveforms of mainshock (black) and EGF (red) waveforms (left) filtered in the frequency range from 0.05 to 0.5Hz (green zone). The numbers at waveform's panels are the maximum amplitudes (in counts). Right panels are the corresponding Fourier amplitude spectra. An example of broadband station ARG.

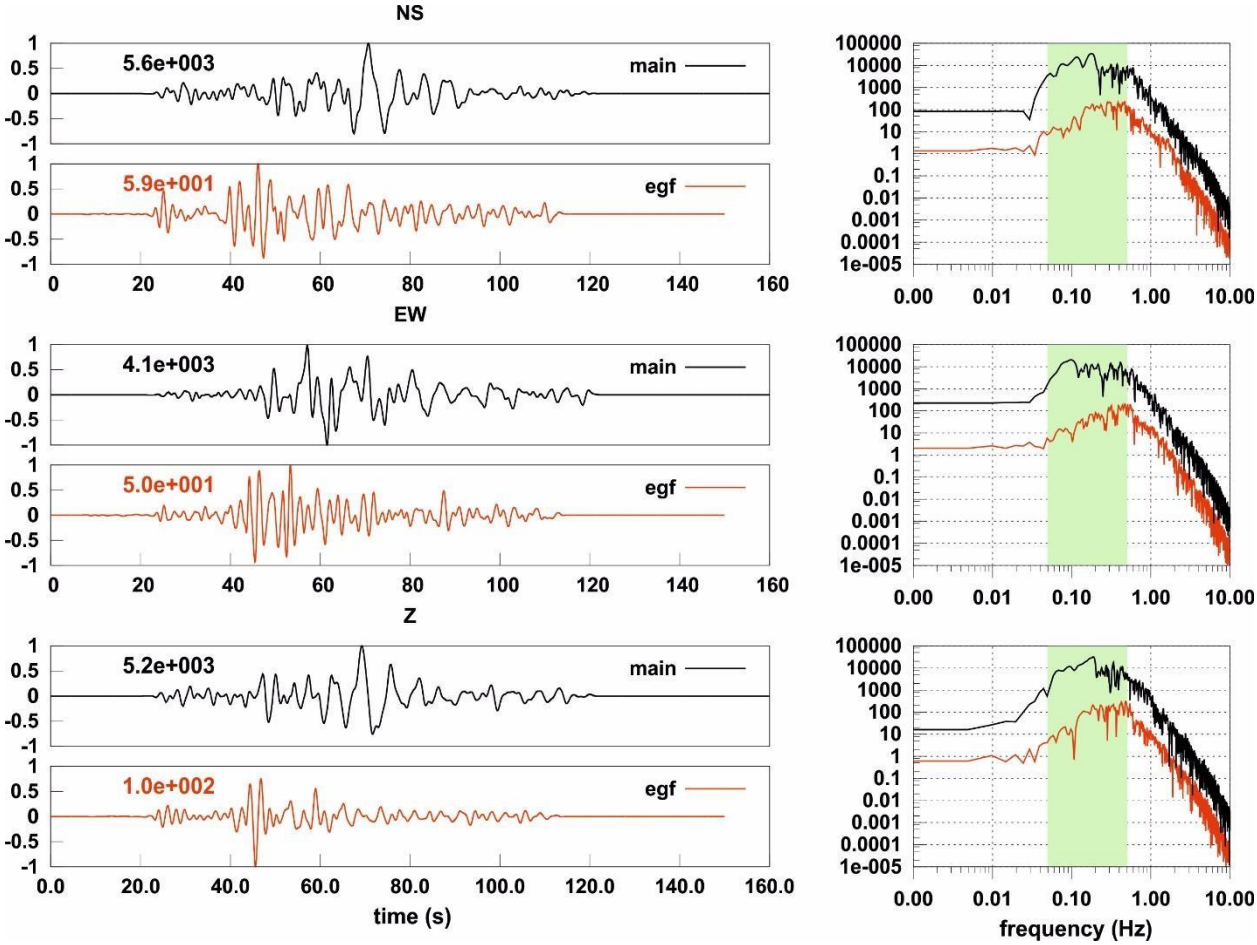


Figure S10. Same as Fig. S9 but for station ASTA.

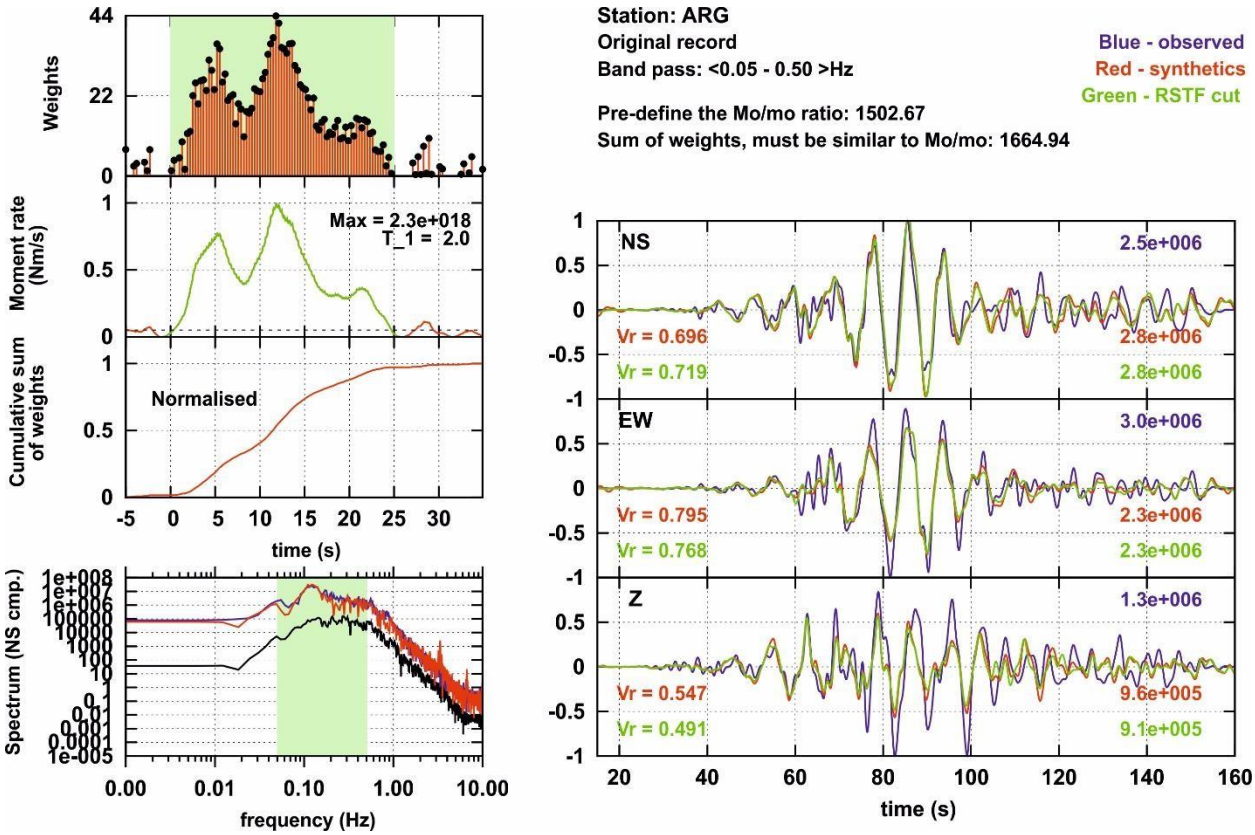


Figure S11. Summary plot of the EGF method for station ARG. See the text for a detailed explanation.

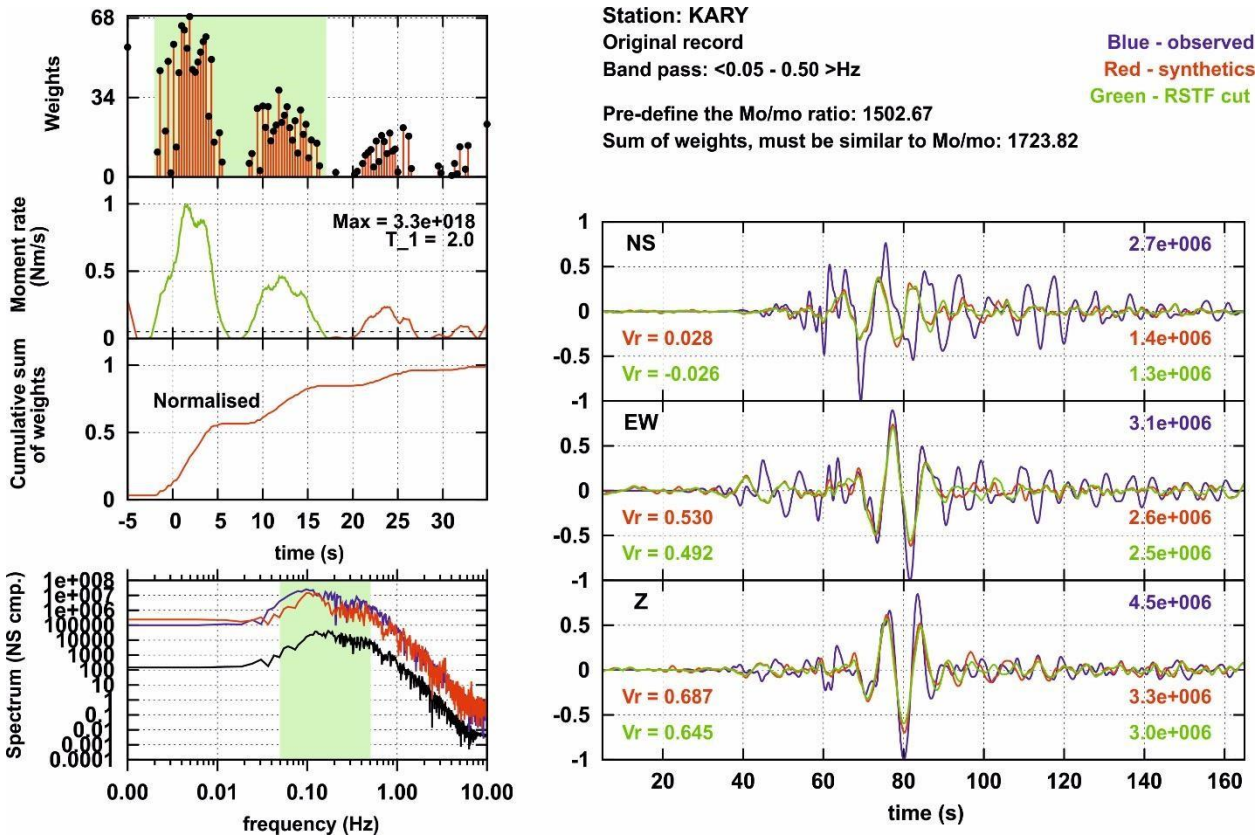


Figure S12. Same as Fig. S11 but for station KARY.