Open Access

Earth System
Science
Data

Enriching the GEOFON seismic catalog with automatic energy magnitude estimations

Dino Bindi[1], Riccardo Zaccarelli[1], Angelo Strollo[1], Domenico Di Giacomo[2], Andres Heinloo[1], Peter Evans[1], Fabrice Cotton[1,3], and Frederik Tilmann[1,4]

[1]German Research Centre for Geoscience GFZ, Potsdam, Germany
[2]International Seismological Center ISC, Thatcham, UK
[3]Institute of Geociences, University of Potsdam, Potsdam, Germany
[4]Institute of Geological Sciences, Freie Universität Berlin, Berlin, Germany

Correspondence: Dino Bindi (bindi@gfz-potsdam.de)

Abstract. We present a seismic catalog (Bindi et al., 2024, https://doi.org/10.5880/GFZ.2.6.2023.010) including energy magnitude M_e estimated from P waves recorded at teleseismic distances in the range $20° \leq \Delta \leq 98°$ and for depths shorter than 80 km. The catalog is built starting from the event catalog disseminated by GEOFON (GEOFOrschungsNetz), considering 6349 earthquakes with moment magnitude $M_w \geq 5$ occurring between 2011 and 2023. Magnitudes are computed using 1 031 396 freely available waveforms archived in EIDA (European Integrated Data Archive) and IRIS (Incorporated Research Institutions for Seismology) repositories, retrieved through the standard International Federation of Digital Seismograph Networks (FDSN) web services (https://www.fdsn.org/webservices/, last access: March 2024). A reduced, high-quality catalog for events with $M_w \geq 5.8$ and from which stations and events with only few recordings were removed forms the basis of a detailed analysis of the residuals of individual station measurements, which are decomposed into station- and event-specific terms and a term accounting for remaining variability. The derived M_e values are compared to M_w computed by GEOFON and with the M_e values calculated by IRIS. Software and tools developed for downloading and processing waveforms for bulk analysis and an add-on for SeisComP for real-time assessment of M_e in a monitoring context are also provided alongside the catalog. The SeisComP add-on has been part of the GEOFON routine processing since December 2021 to compute and disseminate M_e for major events via the existing services.

1 Introduction

Several magnitude scales have been defined to characterize the size of an earthquake. We can, however, divide magnitude scales in two groups, with one including magnitudes based on the amplitudes and periods of different seismic phases measured on band-limited signals (e.g., the body and surface wave magnitudes; Gutenberg, 1945a, b) and the other including magnitude scales related to estimations of macroscopic physical parameters of the earthquake source. The latter comprise the moment (M_w; Kanamori, 1977; Hanks and Kanamori, 1979) and the energy (M_e; Boatwright and Choy,

1986) magnitudes, which are based on seismic moment (Aki, 1966) and radiated seismic energy (Haskell, 1964), respectively. These two magnitude scales are somewhat complementary because, although both represent an estimation of earthquake-related energy, they are determined by different parts of the source spectrum. The seismic moment extrapolated from the low-frequency end and represents the release of elastic energy stored in the Earth's crust or mantle being proportional to the integrated slip across the fault surface. The radiated seismic energy describes the fraction of the total energy released being radiated as seismic waves across all frequencies; i.e., it depends on the earthquake dynamics such

as rupture velocity but also stress drop. M_e estimates have been shown to play an important role when used in conjunction with M_w to better characterize the tsunami and shaking potential of an earthquake (Newman and Okal, 1998; Di Giacomo et al., 2010).

M_w is routinely computed from long period signals of broadband recordings, and it has become a robust and reliable source parameter for large and moderate earthquakes worldwide (Di Giacomo et al., 2021). On the other hand, the computation of M_e is hindered by the necessity of integrating the velocity power spectra over a wide frequency range, while using signals in a limited bandwidth and taking into account propagation effects at high frequencies.

Aiming at validating and testing the procedures for operational purposes, we present a seismic catalog of M_e computed, following the methodology proposed by Di Giacomo et al. (2008) and Di Giacomo et al. (2010) for the rapid assessment of energy magnitude (i.e., without requiring additional source information other than the hypocentral location). The approach is based on the analysis of spectra computed for teleseismic vertical component P waveforms. Teleseismic P waves are commonly used to compute M_e for global earthquakes, as their energy loss during propagation can be more reliably modeled compared to S waves. We further present a detailed analysis of the residuals in a reduced high-quality catalog for events with $M_w \geq 5.8$ with respect to the M_w available in the GEOFON (GEOFOrschungsNetz) catalog and the M_e values computed by IRIS.

2   Energy magnitude computation

2.1   Single-station estimation

We implement the methodology proposed by Di Giacomo et al. (2008) and Di Giacomo et al. (2010) to compute M_e. Teleseismic vertical component P waveforms (BHZ channels, where B stands for broadband, H for high-gain seismometer, and Z for vertical component) are analyzed in the distance range from 20° to 98° and for earthquakes shallower than 80 km. Standard teleseismic range usually starts at 30°, but we use 20° to allow closer stations to be used for rapid response purposes. Shorter distances, however, are difficult to include for global earthquakes, as regional effects are not well accounted for with a 1-D model. Propagation effects are accounted for by frequency-dependent amplitude decay functions that are computed numerically (Wang, 1999) for the ak135Q model (Kennett et al., 1995; Montagner and Kennett, 1996) in the frequency range 0.012–1 Hz.

An estimate of radiated seismic energy E_s is obtained for each single station from the integral of the power spectra of the vertical component P waveform, which is corrected for propagation effects (Haskell, 1964):

$$E_s = \left[\frac{2}{15\pi\rho\alpha^5} + \frac{1}{5\pi\rho\beta^5} \right] \int_{f_1}^{f_2} \left| \frac{\dot{u}(f)}{G(f)/2\pi f} \right|^2 df, \tag{1}$$

where α, β, and ρ are the P-wave velocity, S-wave velocity, and the density at the source, respectively. f is the frequency, and $f_1 = 0.012$ Hz and $f_2 = 1$ Hz are the lower and upper limits of the considered spectral bandwidth. $\dot{u}(f)$ is the P-wave velocity spectrum. $G(f)$ is the median value of Green's function spectrum for displacement, computed from multiple combinations of focal mechanisms and varying strike, dip, and rake over a regular grid (Di Giacomo et al., 2008).

We used analysis windows starting 10 s before the P arrival and with lengths of 90 s for $M_w \leq 7.5$, 120 s for $7.5 < M_w \leq 8.5$, and 180 s for $M_w > 8.5$. The energy magnitude M_e estimate for a single-event station pair is in turn computed as $M_e = 2/3(\log_{10} E_s - 4.4)$, with E_s given in joules (Bormann et al., 2002). The procedure provides M_e estimates at each recording station that can be averaged to minimize path-specific deviations not accounted for by the theoretical model (e.g., directivity and focal mechanism effects and regional variations in attenuation).

2.2   Open-source tool for computing M_e

The above procedure is implemented in the package me-compute (Zaccarelli, 2023). The program uses stream2segment (Zaccarelli et al., 2019; Zaccarelli, 2018) to download events, station metadata, and waveforms from International Federation of Digital Seismograph Networks (FDSN)-compliant repositories in a structured query language (SQL) database.

In our application, the download is configured to fetch events from the GEOFON (Quinteros et al., 2021) event web service, selecting events with computed M_w in the time span 2011–2023. Waveforms are download from EIDA (Strollo et al., 2021) and IRIS (https://service.iris.edu/, last access: March 2024) data centers. The processing routine is implemented in a Python module, which computes the station energy magnitude for each downloaded waveform segment, as summarized in Sect. 2.1, and then calculates the event energy magnitude M_e as the mean of all station magnitudes within the 5th–95th percentile range.

The final output consists of the following files:

- a tabular file in a hierarchical data format (HDF), where each row represents the metadata and measurements, specifically also the station energy magnitude estimate, for a single waveform;

- a tabular file in a comma-separated value (CSV) format aggregating the results of the previous file, where each row represents a seismic event, reporting the event data end metadata and including the M_e estimate for the event;

- an HTML (HyperText Markup Language) file visualizing the selected content reported in the CSV file, where the information for each event can be visualized on an interactive map; and

- one file per processed event in the Quake Markup Language (QuakeML) format, where we also included the M_e value.

All files produced by me-compute are disseminated in the data archive (Bindi et al., 2024; https://doi.org/10.5880/GFZ.2.6.2023.010), along with the stream2segment and me-compute configuration files.

3 Catalog compilation

We use me-compute to compute M_e for $M_w \geq 5$ earthquakes since 2011 in the GEOFON catalog. Table 1 summarizes the steps followed to compile the disseminated M_e catalog. The catalog reports the single-waveform energy magnitude M_{eij} estimated at station j for earthquake i. The energy magnitude M_e for each considered event i is then computed as the median of M_{eij} over the set of recording stations, without considering station static corrections. The starting dataset D0 consists of more than 1 000 000 waveforms (channels BHZ) generated by 6963 earthquakes recorded by 7765 stations belonging to 246 different networks. Only recordings with an average signal-to-noise ratio (SNR) for the amplitude greater than 3 within the frequency range of interest are included in D0. Several integrity and quality checks are applied to remove outliers and faulty signals. Dataset D1 is obtained by analyzing the median residual at the network level, discarding 14 networks characterized by median residuals outside the 2.5 and 97.5 percentile range (Fig. 1a). Dataset D2 is then generated by analyzing the station median values and excluding 382 stations with residuals outside the 2.5–97.5 percentile range (Fig. 1b). Most of the networks and stations removed will have instrumental problems or faulty metadata regarding instrument responses, although in some cases stations with very strong site effects might also be excluded.

An anomaly score is computed to further refine the dataset by flagging anomalous amplitudes using the software sdaas (Zaccarelli, 2022). The software, developed from the work of Zaccarelli et al. (2021), is based on a machine learning algorithm specifically designed for outlier detection (Isolation Forest) which computes an anomaly score in [0, 1], representing the degree of belief of a waveform to be an outlier. The score can be used to assign robustness weights or to define thresholds above which data can be discarded. After inspecting the distribution of the anomaly scores, we set the threshold to 0.62 for $M_w < 7.5$ and to 0.80 for $M_w \geq 7.5$.

The spatial distribution of events and stations generating dataset D3 (Bindi et al., 2024) are shown in Fig. 2a, b. The corresponding M_e residuals are shown in Fig. 3 against distance and M_w. The largest positive residuals correspond mostly to earthquakes with $M_w < 6$ recorded at distances $\Delta > 60°$, where the implemented methodology is expected to generate biased station M_e estimates due to the limitations in the analyzed bandwidth and low signal-to-noise ratio (Di Giacomo et al., 2008, 2010). The overall residual distribution is unbiased and does not show trends of the mean value with distance and magnitude.

Therefore, we further limit the dataset by only considering events with $M_w \geq 5.8$ and at least 10 single-station measurements; we further exclude stations with fewer than 10 recordings in total. We added a column in the disseminated D3 dataset to flag lines corresponding to D6. It consists of $\sim 750\,000$ waveforms for 1671 earthquakes and 7135 stations. The event and station locations of D6 are shown in Fig. 2c and d.

4 Quality assessment via residual analysis

We perform residual analysis to validate the D6 catalog. The relationship between M_e and M_w is analyzed by performing the following mixed-effects regression (Bates et al., 2015; Stafford, 2014):

$$M_{eij} = c_1 + c_2 M_{wi} + \delta S_j + \delta E_i + \epsilon_{ij}, \tag{2}$$

where M_{eij} is the single-waveform energy magnitude estimate at station j for earthquake i, M_{wi} is the moment magnitude of earthquake i, intercept c_1 and slope c_2 parameters define the median model, and δS_i and δE_j are terms that capture station-specific and earthquake-specific adjustments, respectively. ϵ_{ij} accounts for the leftover effects (i.e., residuals that are specific to a particular path/waveform). The random effects δS, δE, and ϵ are zero mean normal distributions by construction. In particular, δS_j (inter-station residual) can represent site effects or instrumental gain corrections, with most of the latter probably removed by the outlier filtering stages described above. The interevent residual δE_i is an event-specific deviation from the M_e expected for a given M_w from the linear regression term. Finally, ϵ_{ij} can be thought of as a noise term for individual measurements, which can be either related to path-specific heterogeneity in attenuation with respect to the 1-D reference model or the influence of ambient noise on the actual measurement.

The interevent and inter-station term distributions are shown in Fig. 4, which are described by standard deviations of $\tau = 0.246$ and $\phi_S = 0.188$ m.u., respectively; the standard deviation of the ϵ is $\phi_0 = 0.232$ m.u. When combining the interevent variability τ with the intra-event variability equal to $\phi = \sqrt{\phi_0^2 + \phi_S^2}$, we obtain the total standard deviation $\sigma = \sqrt{\tau^2 + \phi^2} = 0.407$, which represents the variability in the single-station M_{eij} residuals with respect to the average M_e computed per event. It is worth noting that the δS_j values can be used as station corrections to compute the energy magnitude of new events. In this case, the inter-station contribution

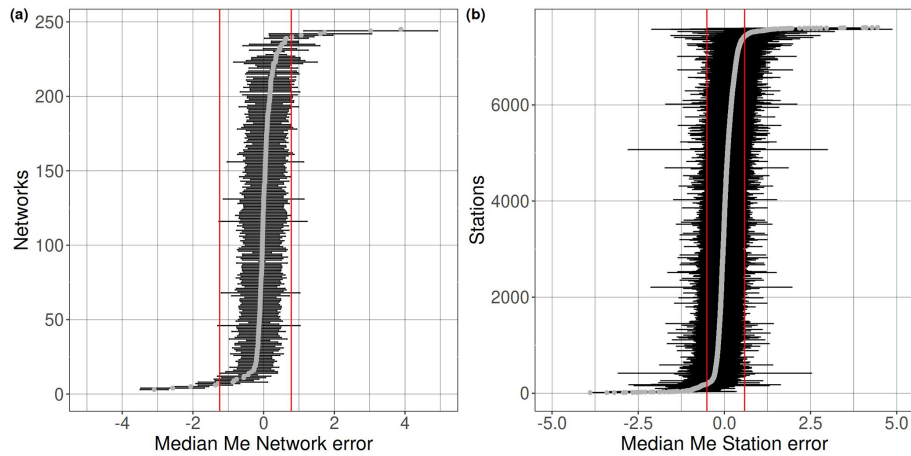**Figure 1.** Median network residuals (circles) for dataset D0 **(a)** and median station residuals for dataset D1 **(b)**. Red lines correspond to the 2.5 and 97.5 percentiles of the distributions. For each network **(a)** and station **(b)**, the horizontal bars correspond to the interval median (circle) ± 1 median absolute deviation (MAD). A few values falling outside the range considered for the horizontal axis are not shown.
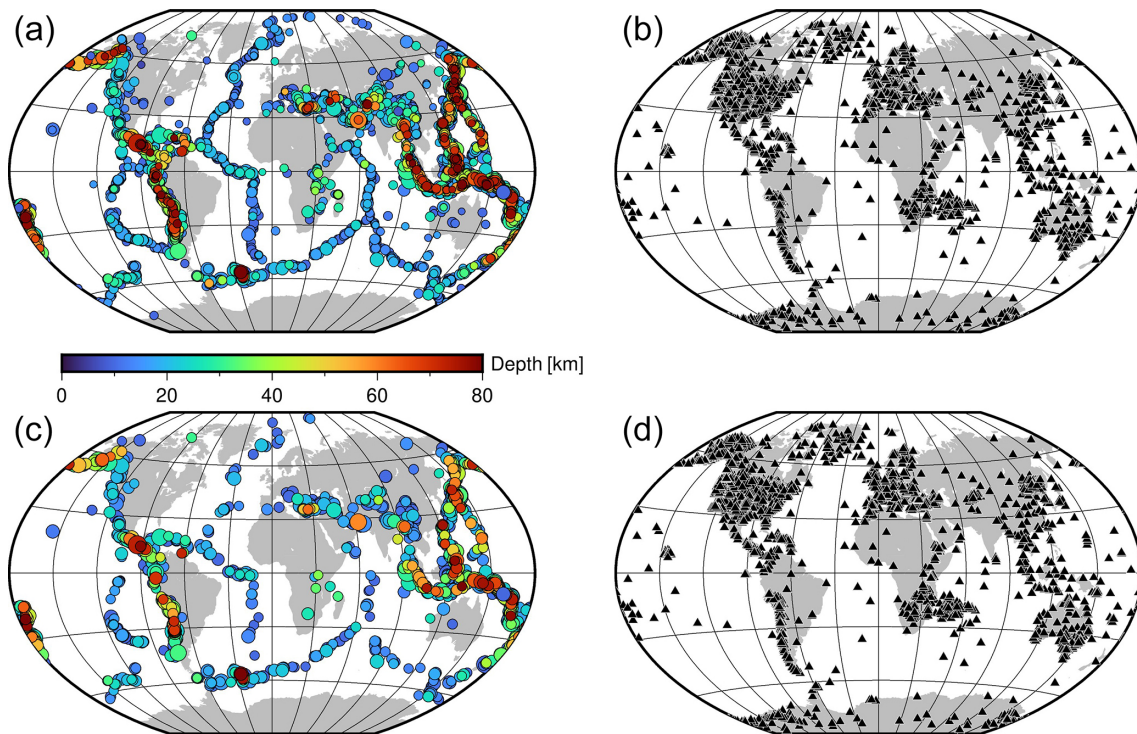

Figure 2. Panels **(a)** and **(b)** show event and station locations for dataset D3 (Table 1), respectively. Panels **(c)** and **(d)** show event and station locations for dataset D6 (Table 1), respectively.

to the total variability is removed, and the expected variability in the M_{eij} distribution is reduced to $\sqrt{\tau^2 + \phi_0^2} = 0.338$. Finally, the linear regression model is defined by the coefficients $c_1 = (0.77 \pm 0.09)$ m.u. and $c_2 = (0.92 \pm 0.01)$. Considering the simplicity of the linear model in Eq. (2) and the large dataset analyzed, the uncertainty in the median model

(sometimes referred to as σ_μ; Atik and Youngs, 2014) is very low, increasing from 0.007 for $M_w = 6$ to 0.039 for $M_w = 9$.

We show the spatial distribution of δS in Fig. 5. Since M_{eij} is computed considering spectral values below 1 Hz and using teleseismic recordings for distances above 20°, δS capture station-specific effects connected to large-scale geological and tectonic crustal features, as exemplified in Fig. 5b for stations located in Europe. Positive δS (i.e., M_{eij} larger than

Table 1. Datasets considered in this study. Datasets in bold are available in Bindi et al. (2024).

Dataset	Records	Networks	Stations	Events	Selections applied sequentially
D0	1 126 465	246	7765	6963	$M_{\mathrm{w}} \geq 5$
D1	1 072 381	232	7617	6944	Network selection (2.5–97.5 %)
D2	1 034 833	228	7235	6880	Station selection (2.5–97.5 %)
D3	1 031 396	228	7234	6349	Anomaly score (< 0.62, < 0.8 for $M_{\mathrm{w}} < 7.5$, ≥ 7.5, resp.)
D4	754 025	228	7228	1731	$M_{\mathrm{w}} \geq 5.8$
D5	751 567	227	7135	1731	No. of records per station ≥ 10
D6	750 903	227	7135	1671	No. of records per event ≥ 10
Dg				153	Comparison between D6 and real time
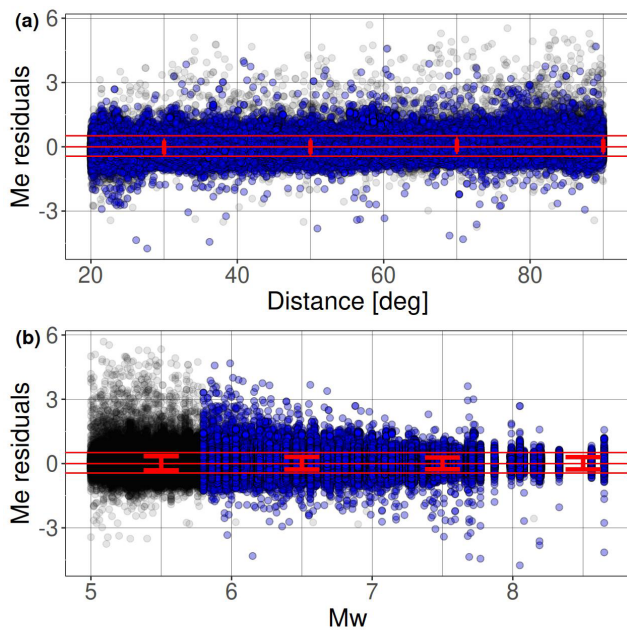
Figure 3. Energy magnitude residuals versus distance **(a)** and moment magnitude **(b)** for dataset D3. Blue dots indicate residuals also included in D6. The horizontal red lines bound the 90 % confidence interval $[-0.43, 0.50]$ of the residual distribution. The error bars indicate the mean ± 1 standard deviation of the residuals computed over different distance (20° wide) and magnitude (1 m.u. wide) intervals.

the median) are observed for stations located in basins like in the Po plain, in the Moesian region, in the Netherlands, and in the East Anatolian Fault region. Negative values δS (i.e., $M_{\mathrm{e}ij}$ lower than the median) are observed for stations located in mountain ranges such as the Pyrenees, the Alps, or in Harz Highlands and also tectonically highly active regions that are as cratonic as the East African Rift. The station terms can represent both site amplification, e.g., for stations in sedimentary basins, and anomalously high or low attenuation in the crust and or mantle surrounding the station. The station-specific residuals are disseminated along with the catalog to allow the computation of M_{e} for future earthquakes,

while taking into account static magnitude corrections to reduce variability.

The spatial distribution of the interevent variability, δE, is shown in Fig. 6 for the smallest and largest values.

Considering depths shallower than 30 km (Fig. 6a and b), continental Asia, the Philippines, and Indonesia, as well as the Aleutian Islands, show positive values. California, Mexico, central America, and the Mid-Atlantic Ridge are characterized mostly by negative values. Considering deeper events (Fig. 6c and d), Japan and the Philippines have mostly positive values, while Mexico and central America have mostly negative values. The event-specific residuals are also disseminated, along with the catalog to increase the usefulness of the product from the event point of view, to allow the user to perform further refinements.

Path-specific residuals ϵ are shown in Fig. 7 for three selected receiving areas in Europe, California, and Australia. Since in the partition of the residuals the leftover distribution ϵ represents the component not related to systematic station and event effects, they are mostly connected to lateral variability in attenuation in the Earth's interior with respect to the used global 1-D model and amplitude variation related to P-wave radiation patterns for different focal mechanisms.

Finally, the $M_{\mathrm{e}ij}$ versus M_{w} scaling defined by the linear regression coefficients c_1 and c_2 of Eq. (2) is shown in Fig. 8.

4.1 Catalog validation: comparison with IRIS

The energy magnitude computed in this study is compared to the values disseminated by IRIS through the SPUD service IRIS DMC (2013). The methodology implemented by IRIS is described by Convers and Newman (2011) and based on the analysis of Boatwright and Choy (1986) and Newman and Okal (1998). Similar to our approach, the energy flux is computed from the P-wave group (P + pP + sP) in the frequency domain. The single-station estimations are corrected for frequency-dependent anelastic attenuation effects and converted back to the energy radiated by the source by applying corrections for geometrical spreading, depth, and mechanism-dependent effects for P waves and considering a theoretical partition of the energy between P and S waves. The energy is computed considering the frequency
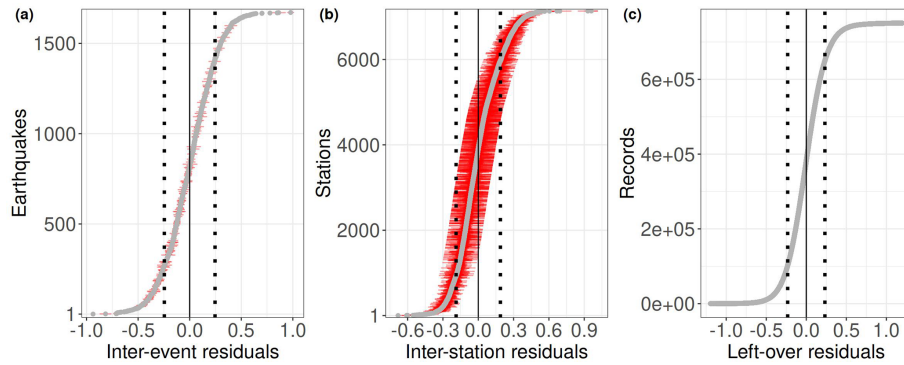
Figure 4. Cumulative distribution functions for event δE **(a)**, station δS **(b)**, and leftover ϵ distributions (circles) determined according to the mixed-effects regression in Eq. (2) applied to dataset D6. Dotted lines correspond to standard deviations $\pm 1\tau$ **(a)**, $\pm 1\phi_S$ **(b)**, and $\pm 1\phi_0$ **(c)**. Horizontal red lines in panels **(a)** and **(b)** are the standard errors in the random effects. In panel **(c)**, values of ϵ exceeding ± 1.2 in absolute value are not shown.
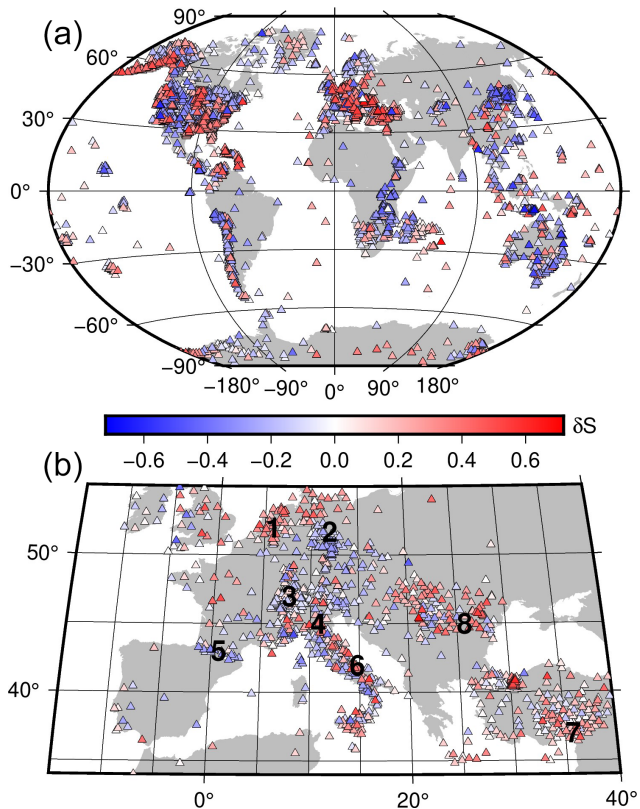
Figure 5. (a) Distribution of the site-specific residuals δS (see Eq. 2) and **(b)** magnified view over a portion of Europe. Numbers in panel **(b)** indicate the following locations: (1) Netherlands; (2) Harz highlands, Germany; (3) Switzerland; (4) Po plain, Italy; (5) Pyrenees mountain range; (6) Apennine Mountains; (7) East Anatolian Fault region; and (8) Moesian Platform.

range 0.014–2 Hz (broadband for M_e(BB)) or 0.5-02 Hz (high frequency for M_e(HF)), analyzing stations in the distance range 25–80°. The duration of the time window used for the computation is based on analysis of the cumulative high-frequency energy (0.5–2 Hz) as a function of time. The crossover time used to compute the energy flux is identified at the intersection between the near-constant increasing rate for short times and the relative flat asymptotic behavior for long durations. The SPUD service disseminates both the high-frequency M_e(HF) and broadband M_e(BB) estimates.

Two regression models are calibrated against the broadband and high-frequency estimates disseminated by IRIS through SPUD. The best-fit models, shown in Fig. 9, are $M_e = (-0.076 \pm 0.229) + (1.002 \pm 0.033) M_e(\text{HF})$ and $M_e = (0.795 \pm 0.188) + (0.896 \pm 0.027) M_e(\text{BB})$, with the standard deviation of the residuals equal to 0.234 and 0.175, respectively. For the magnitude range from 6 to 8, this results in biases of 0.06 m.u. for M_e vs. M_e(HF) and varying from 0.17 to -0.04 m.u. for M_e vs. M_e(BB); i.e., our estimates are nearly unbiased relative to M_e(HF) and tend to slightly overestimate M_e(BB) at the lower end of the applicability range.

4.2 Catalog validation: role of style of faulting

The faulting style is classified into normal, reverse, and strike slip categories, based on the plunge of the P, T, and N axes (Frohlich and Apperson, 1992), as extracted from the GEOFON moment tensor solutions. There is a normal fault (NF) if the plunge $(P) \geq 60°$, a strike slip (SS) if the plunge $(N) \geq 60°$, and a thrust fault (TF) if the plunge $T \geq 50°$. In the other cases, the earthquake is labeled with OF (other faulting styles). To investigate the role of the style of faulting (SOF), we separate the event term into a fixed offset for each SOF class and a perturbation term for each event. If we indicate the classes of the SOF grouping factor (corresponding to NF, SS, TF, and OF) with $k = 1, 2, 3, 4$ and the class of event i with k_i, then the equation for the extended mixed-effects
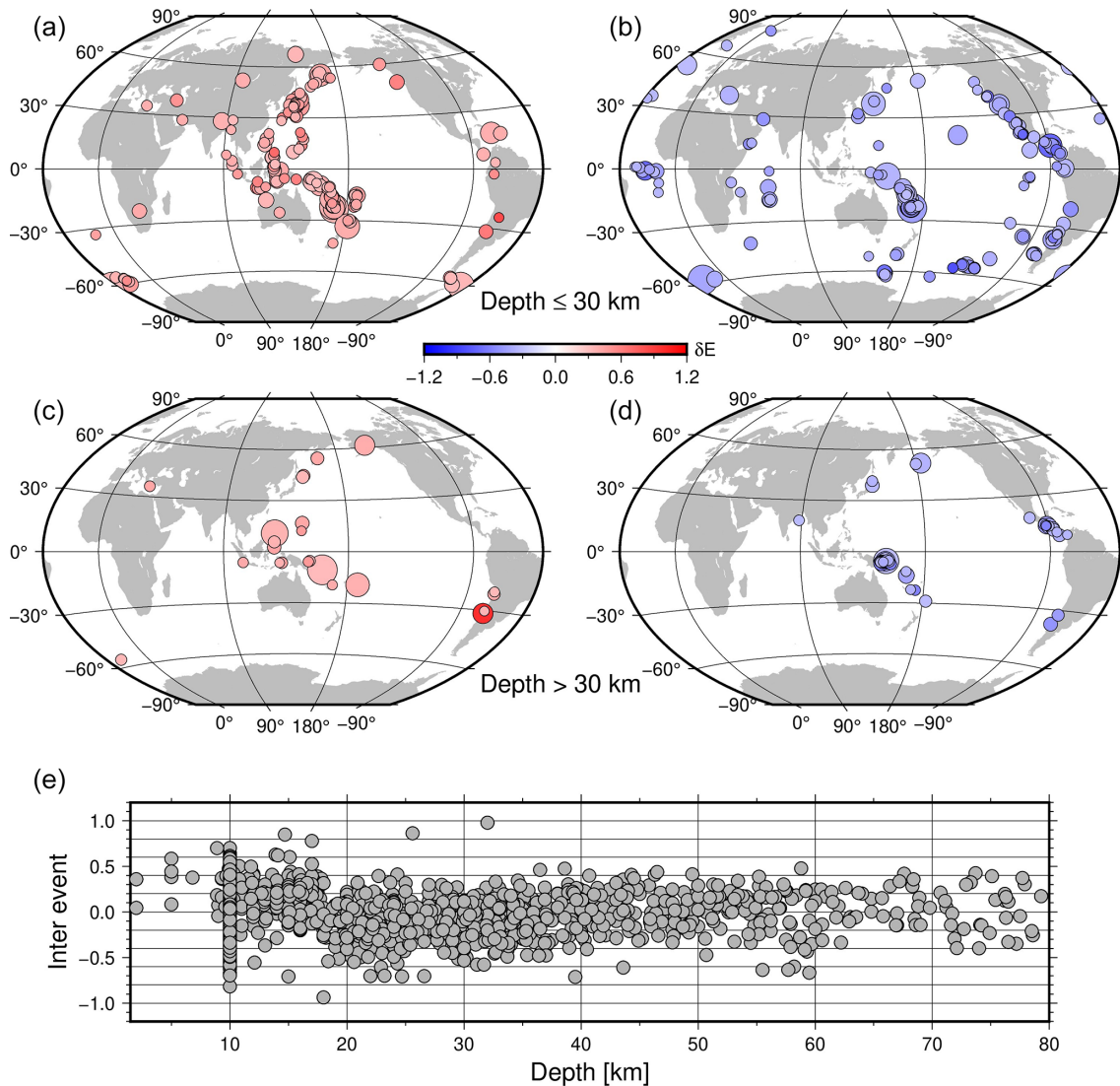
Figure 6. Extreme values for event specific residuals δE for the M_{eij} versus M_w mixed-effects model of Eq. (2). Only values below the 10th percentile (panels **b** and **d**) and above the 90th percentile (panels **a** and **c**) of the distribution are shown (the percentiles are about ± 0.3). In panels **(a)** and **(b)**, earthquakes with hypocentral depths shallower than 30 km are selected. In panels **(c)** and **(d)**, events deeper than 30 km are considered. The distribution of δE versus depth for all events is shown in panel **(e)**.

model is

$$M_{eij} = e_1 + e_2 M_{wi} + \delta S_j + [\delta\text{SOF}_{k_i} + \delta E_{\text{SOF}_i}] + \epsilon_{ij}, \quad (3)$$

where δSOF are the terms characterizing the average effects of the different SOF and δE_{SOF} are accounting for interevent differences within each SOF class (nested random effects). The standard deviations of the δS, δSOF, δE_{SOF}, and ϵ distributions are $\phi_S = 0.190$, $\tau_{\text{SOF}} = 0.095$, $\tau = 0.236$, and $\phi_0 = 0.232$, respectively, generating a total standard deviation $\sigma = 0.393$. The SOF terms are $\delta\text{SOF}_1 = 0.098$ (NF), $\delta\text{SOF}_2 = -0.108$ (SS), $\delta\text{SOF}_3 = -0.045$ (TF), and $\delta\text{SOF}_4 = 0.055$ (OF) (Fig. 10). The largest difference is between SS and NF, a total of 0.206 m.u. There is a systematic impact of the SOF on the intercept of the model, but the associated variability

is smaller compared to the interevent variability τ (in other words, SOF effects are statistically significant, but the distributions of interevent terms separated according to faulting style are strongly overlapping).

The SOF effects might arise due to physical differences (on average) between the different faulting types, e.g., due to systematically different stress drops, differences in the maturity of faults or typical environments (intra-plate vs. interplate), where different faulting types occur most often, or they might be artifacts due to the fact that the (Di Giacomo et al., 2008) method used here does not account for radiation pattern effects, and the teleseismic arrivals utilized here preferentially sample certain parts of the focal sphere. Therefore, we also investigate the role of the SOF in the relationship be-
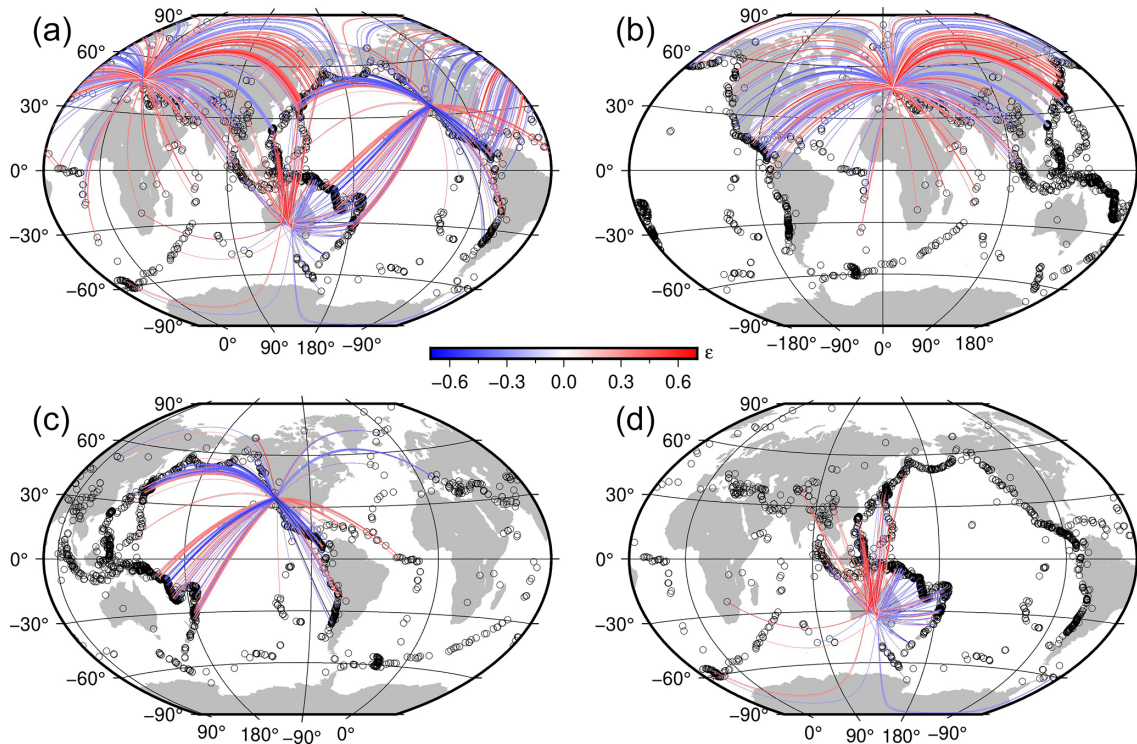
Figure 7. (a) Residual distribution ϵ of Eq. (2) for three different receiving areas, showing only $|\epsilon| > 0.30$. **(b)** As in panel **(a)** but considering only the European receiving area. **(c)** As in panel **(a)** but considering only the receiving area in California. **(d)** As in panel **(a)** but considering only the receiving area in Australia. Circles indicate the earthquake locations.
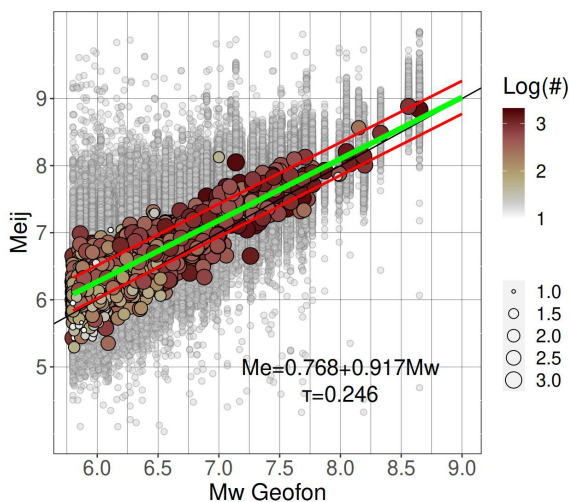

Figure 8. M_{eij} versus M_w scaling. Gray circles are the station M_{eij} estimates, and filled circles represent event M_e values calculated as medians of all station estimates for that event. Color indicates how many stations contributed to each estimate. The best-fit line in green is derived from the mixed-effects regression (Eq. 2), considering ± 1 interevent standard deviation τ (red lines). The faint black line shows equality for reference.

tween M_e derived in this study and the $M_e(\text{HF})$ and $M_e(\text{BB})$ values disseminated by IRIS. We recall that the methodology implemented by IRIS accounts for radiation pattern effects, which are related to the SOF. For this analysis, the regression model is the following:

$$M_e = g_1 + g_2 M_{\text{iris}} + \delta \text{SOF} + \epsilon, \tag{4}$$

where M_{iris} is either $M_e(\text{HF})$ or $M_e(\text{BB})$. Results shown in Fig. 11 confirm that the largest intercept difference is between normal and strike slip events, and the differences in terms of magnitude units (m.u.) are also similar between the other SOF. This suggests that a large part of the SOF term is influenced by radiation pattern effects, and interpretations of these differences in terms of geodynamics or hazard potential should be done very cautiously.

5   Real-time module for SeisComP

The module, derived from me-compute has been integrated to the SeisComP package (Helmholtz Centre Potsdam GFZ German Research Centre for Geosciences and GEMPA GmbH, 2008) and has been part of the GEOFON routine real-time processing since December 2021. The first event for which M_e calculations are availa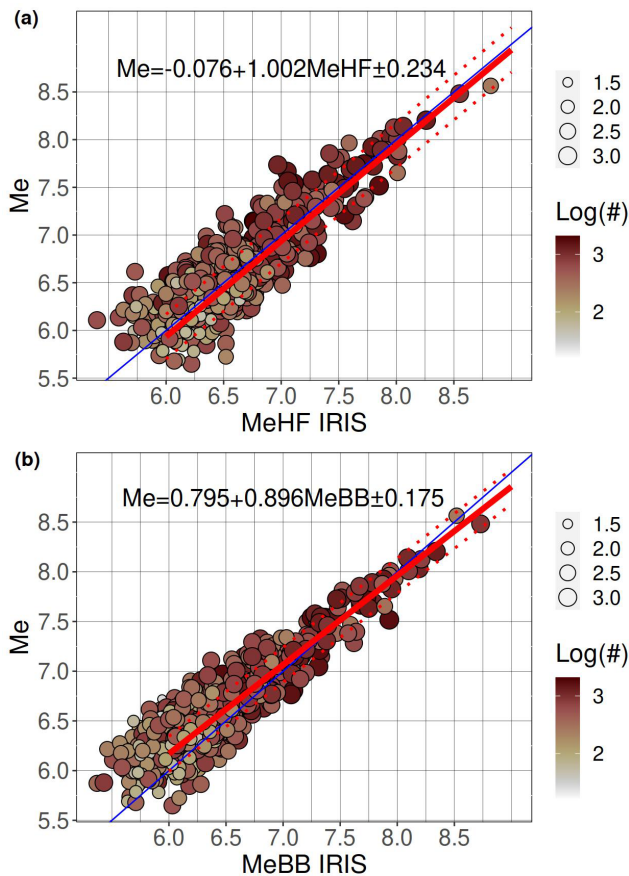ble and disseminated via the usual GEOFON services is https://geofon.gfz-potsdam.de/eqinfo/event.php?id=gfz2021xxzt (last access:

Figure 9. Comparison with energy magnitude disseminated by IRIS considering **(a)** M_e(HF) and **(b)** M_e(BB) (717 common events). The red line shows the linear regression fit, and the dotted lines show 1 standard deviation of the M_e residuals. The blue line shows line of equality for reference.

Figure 10. M_e versus M_w categorized with SOF.

March 2024), which occurred on 7 December 2021 10:28:00.3 UTC (M_e 5.7 and M_w 5.5). The scmert add-on is available at https://github.com/SeisComP/scmert (last access: March 2024).

The add-on has been configured in GEOFON to trigger the calculation for each origin created by the automatic processing with magnitude ≥ 5.5 and to compute station magnitudes M_{eij} for all stations according to the definition of M_e in the distance 20–98°. The scmert procedure is applied with the settings used by the GEOFON Earthquake Information Service, using stations available in real time from the GEOFON Extended Virtual Network (https://geofon.gfz-potsdam.de/eqinfo/gevn/, last access: March 2024), including station selection and the distribution trimming of 25 %. The workflow for M_e computations is as follows: as soon as an automatically detected event reaches the magnitude threshold, scmert is triggered and starts to compute M_{eij} upon receiving data from stations beyond 20°. The process continues until the selected window length (determined by the actual preliminary magnitude) of the last station at 98° is acquired. The first esti-
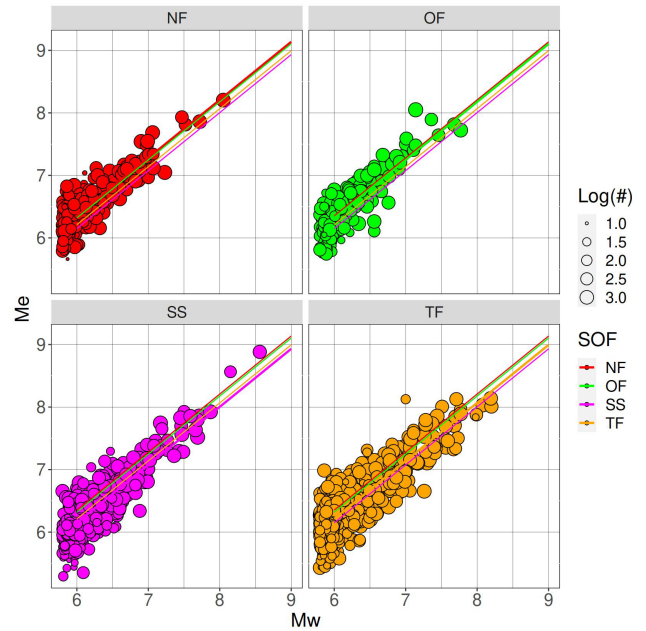
mate of the magnitude M_e is released shortly after collecting 20 M_{eij} estimates from individual station, usually within a few minutes of the earthquake's origin time. SeisComP modules continue to refine the estimate until no further updates are required (this includes manual release at later stages). The computed station magnitudes M_{eij} are fully integrated also into the SeisComP Origin Locator View Graphical User Interface (scolv GUI; Fig. 12), with station magnitudes and residuals displayed in a dedicated energy magnitude tab.

The energy magnitude values from both modules are compared in Fig. 13. We used scmert with the same settings as the GEOFON earthquake monitoring service, including station selection and trimming of the distributions. The values are in good agreement, and the best-fit model is $M_e = 0.057 + 0.987 M_e$(GEO), with a standard deviation of 0.118. The average difference computed for magnitudes between 6 and 8 is -0.028.

All values for M_e that have been calculated since the start of the routine processing with scmert can be accessed via the fdsnws-event web service running at GEOFON by specifying M_e as the magnitude type (i.e., https://geofon.gfz-potsdam.de/fdsnws/event/1/query?starttime=2021-12-07&magnitudetype=Me&includeallmagnitudes=true&nodata=404, last access: March 2024). These values are also disseminated to other agencies (e.g., ISC and EMSC) via the usual downstream channels, including a real-time push service.
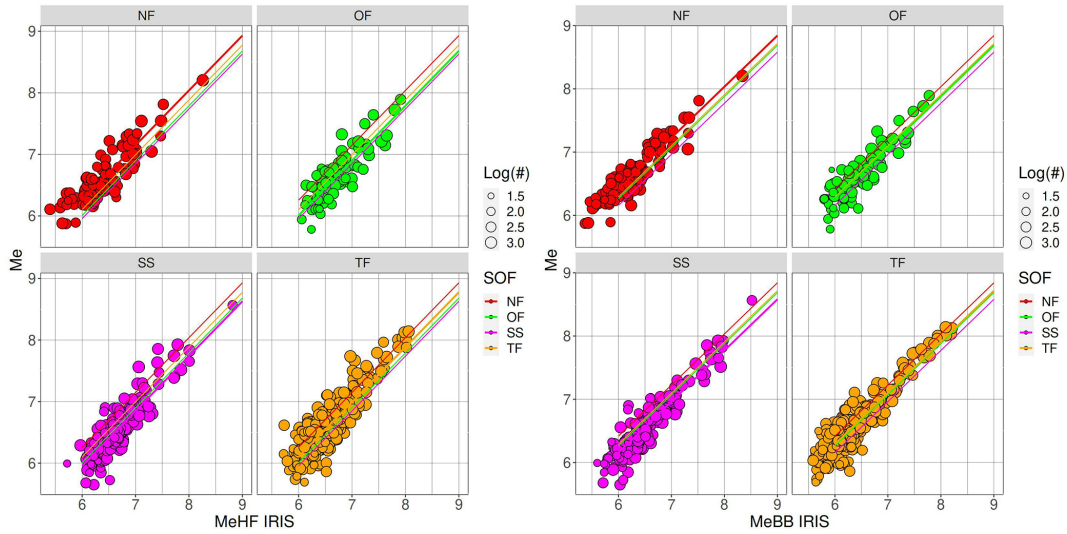
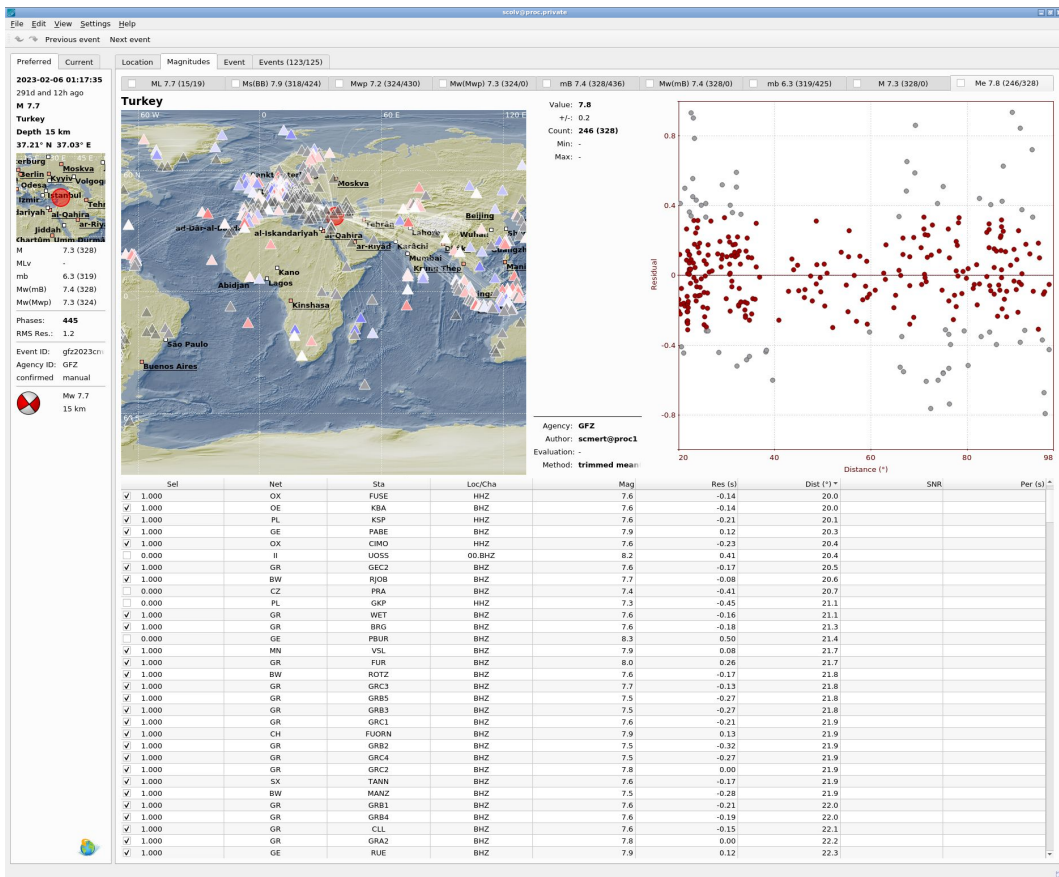**Figure 11.** M_e versus M_e(BB) and M_e(HF), as categorized with SOF.

Figure 12. Screenshot of the SeisComP Origin Locator View (scolv) interactive tool for the M_w 7.7 Republic of Türkiye earthquake that occurred on 6 February 2023, 01:17 UTC along the East Anatolian Fault. The obtained network magnitude value of M_e is 7.8. Stations used are color-coded according to M_e magnitude residuals (top-left frame). Stations in gray are excluded from the network magnitude, do not match the distance range definition, or were trimmed while computing the average magnitude because they are within the $\pm 12.5\%$ range. The top-right scatterplot shows M_e residuals by distance (in red those that contributed to actual M_e network magnitude). The topography shown in the map is generated using the ETOPO1 global relief model (Amante and Eakins, 2009).
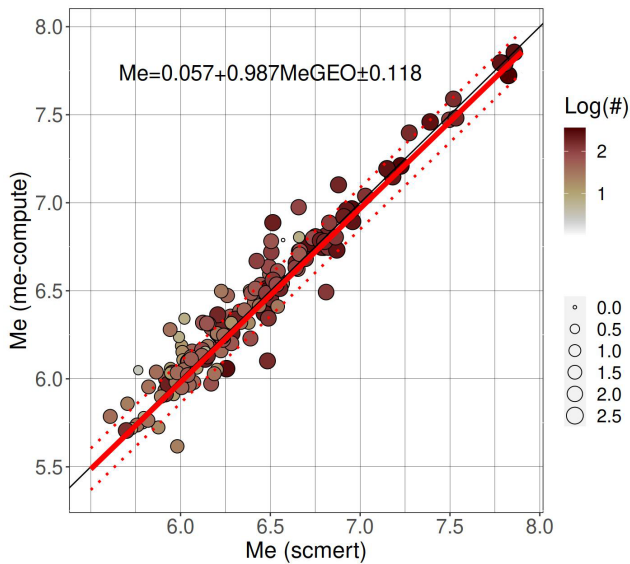
Figure 13. Comparison between M_e computed in real time by GE-OFON, with the scmert add-on for SeisComP (x axis) and offline estimation using me-compute (y axis), considering 153 common events.

6   Code and data availability

Code used for computing the energy magnitude is available from

- offline computations in me-compute https://doi.org/10.5880/GFZ.2.6.2023.008 (Zaccarelli, 2023); and

- real-time computations in SeisComP with scmert https://doi.org/10.5880/GFZ.2.4.2020.003 (Helmholtz Centre Potsdam GFZ German Research Centre for Geosciences and GEMPA GmbH, 2008).

Analyses have been performed in R (https://www.R-project.org/, R Core Team, 2020), and we used the Generic Mapping Tools (https://www.generic-mapping-tools.org/, last access: March 2024; Wessel et al., 2013) to produce Figs. 2, 5, 6, and 7. The archive, including the energy magnitude catalog (D3 and D6 in Table 1), the complete list of references for the seismic networks analyzed in this article with me-compute, and an example of the configuration files, is available from Bindi et al. (2024) (https://doi.org/10.5880/GFZ.2.6.2023.010).

7   Conclusive remarks

We computed the energy magnitude M_e for 6349 events in the moment magnitude catalog disseminated by GEOFON. When combined with M_w, M_e allows a better characterization of the tsunami and shaking potential of an earthquake. The procedure used to compile the dataset, which includes 1 031 396 M_e values for each recording station, is described

in detail. Residuals are evaluated using a mixed-effects regression, which partitions the overall residuals into event-specific and station-specific contributions. These random effects are included in the distributed catalog, enabling the computation of M_e for future events and using inter-station residuals as station corrections to reduce the uncertainty in M_e. They also enable the assessment of energy magnitude adjustments for specific regions or faulting mechanisms, using interevent residuals and locating propagation anomalies with respect to the global model used to compute Green's functions using the leftover residuals. The methodology employed for computing M_e (Di Giacomo et al., 2008) is suitable for the rapid assessment of M_e (Di Giacomo et al., 2010). Therefore, it has been implemented as a module for SeisComP, allowing the automatic computation of M_e in real time and keeping the M_e catalog up to date.

References

Aki, K.: Generation and Propagation of G Waves from the Niigata Earthquake of June 16, 1964. Part 2. Estimation of earthquake moment, released energy, and stress-strain drop from the G wave spectrum, B. Earthq. Res. I. Tokyo, 44, 73–88, https://doi.org/10.15083/0000033586, 1966.

Amante, C. and Eakins, B. W.: ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis, NOAA Technical Memorandum NESDIS NGDC-24. National Geophysical Data Center, NOAA [data set], https://repository.library.noaa.gov/view/noaa/1163 (last access: March 2024), 2009.

Atik, L. A. and Youngs, R. R.: Epistemic Uncertainty for NGA-West2 Models, Earthq. Spectra, 30, 1301–1318, https://doi.org/10.1193/062813EQS173M, 2014.

Bates, D., Mächler, M., Bolker, B., and Walker, S.: Fitting Linear Mixed-Effects Models Using lme4, J. Stat. Softw., 67, 1–48, https://doi.org/10.18637/jss.v067.i01, 2015.

Bindi, D,. Zaccarelli, R., Strollo, A., Di Giacomo, D., Heinloo, A., Evans, P., Cotton, F., and Tilmann, F.: Global energy magnitude catalog 2011–2023 with event selection driven by Mw Geofon, V. 1.1. GFZ Data Services [data set], https://doi.org/10.5880/GFZ.2.6.2023.010, 2024.

Boatwright, J. and Choy, G. L.: Teleseismic estimates of the energy radiated by shallow earthquakes, J. Geophys. Res.-Sol. Ea., 91, 2095–2112, https://doi.org/10.1029/JB091iB02p02095, 1986.

Bormann, P., Baumbach, M., Bock, G., Grosser, H., Choy, G., and Boatwright, J.: Seismic Sources and Source Parameters, in: IASPEI New Manual of Seismological Observatory Practice, edited by Bormann, P., vol. 1, chap. 3, 94 pp., Deutsches GeoForschungsZentrum GFZ, Potsdam, https://gfzpublic.gfz-potsdam.de/pubman/item/item_4015 (last access: March 2024), 2002.

Convers, J. A. and Newman, A. V.: Global Evaluation of Large Earthquake Energy from 1997 Through mid-2010, J. Geophys. Res., 116, B08304, https://doi.org/10.1029/2010JB007928, 2011.

Di Giacomo, D., Grosser, H., Parolai, S., Bormann, P., and Wang, R.: Rapid determination of Me for strong to great shallow earthquakes, Geophys. Res. Lett., 35, L10308, https://doi.org/10.1029/2008GL033505, 2008.

Di Giacomo, D., Parolai, S., Bormann, P., Grosser, H., Saul, J., Wang, R., and Zschau, J.: Suitability of rapid energy magnitude determinations for emergency response purposes, Geophys. J. Int., 180, 361–374, https://doi.org/10.1111/j.1365-246X.2009.04416.x, 2010.

Di Giacomo, D., Harris, J., and Storchak, D. A.: Complementing regional moment magnitudes to GCMT: a perspective from the rebuilt International Seismological Centre Bulletin, Earth Syst. Sci. Data, 13, 1957–1985, https://doi.org/10.5194/essd-13-1957-2021, 2021.

Frohlich, C. and Apperson, K. D.: Earthquake focal mechanisms, moment tensors, and the consistency of seismic activity near plate boundaries, Tectonics, 11, 279–296, https://doi.org/10.1029/91TC02888, 1992.

Gutenberg, B.: Amplitudes of surface waves and magnitudes of shallow earthquakes, B. Seismol. Soc. Am., 35, 3–12, 1945a.

Gutenberg, B.: Amplitudes of P, PP, and S and magnitude of shallow earthquakes, B. Seismol. Soc. Am., 35, 57–69, https://doi.org/10.1785/BSSA0350020057, 1945b.

Hanks, T. C. and Kanamori, H.: A moment magnitude scale, J. Geophys. Res., 84, 2348–2350, https://doi.org/10.1029/JB084iB05p02348, 1979.

Haskell, N. A.: Total energy and energy spectral density of elastic wave radiation from propagating faults, B. Seismol. Soc. Am., 54, 1811–1841, https://doi.org/10.1785/bssa05406a1811, 1964.

Helmholtz Centre Potsdam GFZ German Research Centre for Geosciences and GEMPA GmbH: The SeisComP seismological software package, GFZ Data Services [code], https://doi.org/10.5880/GFZ.2.4.2020.003, 2008.

IRIS DMC: Data Services Products: EQEnergy Earthquake energy & rupture duration, https://doi.org/10.17611/DP/EQE.1, 2013.

Kanamori, H.: The energy release in great earthquakes, J. Geophys. Res., 82, 2981–2987, https://doi.org/10.1029/JB082i020p02981, 1977.

Kennett, B. L. N., Engdahl, E. R., and Buland, R.: Constraints on seismic velocities in the Earth from traveltimes, Geophys. J. Int., 122, 108–124, https://doi.org/10.1111/j.1365-246x.1995.tb03540.x, 1995.

Montagner, J.-P. and Kennett, B.: How to reconcile body-wave and normal-mode reference Earth models?, Geophys. J. Int., 125, 229–248, 1996.

Newman, A. V. and Okal, E. A.: Teleseismic estimates of radiated seismic energy: The E/M_0 discriminant for tsunami earthquakes, J. Geophys. Res.-Sol. Ea., 103, 26885–26898, https://doi.org/10.1029/98JB02236, 1998.

Quinteros, J., Strollo, A., Evans, P. L., Hanka, W., Heinloo, A., Hemmleb, S., Hillmann, L., Jaeckel, K., Kind, R., Saul, J., Zieke, T., and Tilmann, F.: The GEOFON Program in 2020, Seismol. Res. Lett., 92, 1610–1622, https://doi.org/10.1785/0220200415, 2021.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/ (last access: March 2024), 2020.

Stafford, P. J.: Crossed and Nested Mixed-Effects Approaches for Enhanced Model Development and Removal of the Ergodic Assumption in Empirical Ground-Motion Models, B. Seismol. Soc. Am., 104, 702–719, https://doi.org/10.1785/0120130145, 2014.

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, Seismol. Res. Lett., 92, 1788–1795, https://doi.org/10.1785/0220200413, 2021.

Wang, R.: A simple orthonormalization method for stable and efficient computation of Green's functions, B. Seismol. Soc. Am., 89, 733–741, https://doi.org/10.1785/BSSA0890030733, 1999.

Wessel, P., Smith, W. H. F., Scharroo, R., Luis, J., and Wobbe, F.: Generic Mapping Tools: Improved Version Released, Eos, 94, 409–410, https://doi.org/10.1002/2013EO450001, 2013.

Zaccarelli, R.: Stream2segment: a tool to download, process and visualize event-based seismic waveform data (Version 2.7.3), GFZ Data Services [code], https://doi.org/10.5880/GFZ.2.4.2019.002, 2018.

Zaccarelli, R.: sdaas – a Python tool computing an amplitude anomaly score of seismic data and metadata using simple machine-Learning models, GFZ Data Services [code], https://doi.org/10.5880/GFZ.2.6.2023.009, 2022.

Zaccarelli, R.: me-compute: a Python software to download events and data from FDSN web services and compute their energy magnitude (M_e), GFZ Data Services [code], https://doi.org/10.5880/GFZ.2.6.2023.008, 2023.

Zaccarelli, R., Bindi, D., Strollo, A., Quinteros, J., and Cotton, F.: Stream2segment: An Open-Source Tool for Downloading, Processing, and Visualizing Massive Event-Based Seismic Waveform Datasets, Seismol. Res. Lett., 90, 2028–2038, https://doi.org/10.1785/0220180314, 2019.

Zaccarelli, R., Bindi, D., and Strollo, A.: Anomaly Detection in Seismic Data – Metadata Using Simple Machine-Learning Models, Seismol. Res. Lett., 92, 2627–2639, https://doi.org/10.1785/0220200339, 2021.

https://doi.org/10.5194/essd-16-1733-2024

Earth Syst. Sci. Data, 16, 1733–1745, 2024