Data Mine

ITACAext 2.0: A High-Quality Parametric Table of Strong-to-Weak Motion Recordings for Seismological and Engineering Research in Italy

Giovanni Lanzano*[1], Lorenzo Vitrano[1], Dino Bindi[2], Chiara Felicetta[1], Sara Sgobba[1], Lucia Luzi[1], and Francesca Pacor[1]

Abstract

In the most recent release of the ITalian ACcelerometric Archive (ITACA) portal (v. 4.0; Felicetta *et al.*, 2023), a parametric table (intensity measures and associated metadata) of seismic records of interest for the characterization of seismic motion in Italy was made public. This flatfile is called ITACAext and includes accelerometer and velocimeter recordings of $M_w \geq 3.0$ earthquakes and seismic networks in Italy and neighboring countries. The dataset contains parameters of recordings manually processed by experts from 1972 to 2022, with a total of 52,706 observations from 2,338 events and 2,294 stations. This dataset provides the most up-to-date overview of strong motion recordings available for Italy, including the strongest earthquake recordings of all Italian seismic sequences over the last 50 yr (1976 Friuli, 1980 Irpinia, 1997 Umbria–Marche, 2009 L'Aquila, 2012 Emilia, and 2016–2017 central Italy). It also includes a significant number of near-source records of earthquakes in volcanic and subduction areas. The flatfile contains not only the peak amplitudes (peak ground acceleration, peak ground velocity, and peak ground displacement), energy parameters (cumulative absolute velocity, Arias and Housner intensity), and ordinates of the acceleration response spectrum (damping 5%, period interval 0.01–10 s), but also the ordinates of the Fourier spectrum (f = 0.04–50 Hz), calculated for the entire signal as well as for the *S* phase. ITACAext also includes several metadata, all conveniently referenced, including those most relevant to the calibration and testing of ground-motion models (GMMs), such as various magnitude estimates, different distance metrics, focal mechanisms and site characterization parameters such as $V_{S,30}$. To support the analysis of ground motion of tectonic origin at the national scale, a selection of shallow active crustal recordings was performed. In addition, an analysis of the quality of the dataset was carried out via residual analysis in two stages. These checks also consisted of comparing the seismic responses of collocated sensors and support the inclusion of velocimetric recordings in ground-motion analyses. In the end, we flagged 42,877 records from 2,045 events that can be confidently used to calibrate and test active shallow crustal GMMs at the Italian scale.

Cite this article as Lanzano, G., L. Vitrano, D. Bindi, C. Felicetta, S. Sgobba, L. Luzi, and F. Pacor (2025). ITACAext 2.0: A High-Quality Parametric Table of Strong-to-Weak Motion Recordings for Seismological and Engineering Research in Italy, *Seismol. Res. Lett.* **XX**, 1–22, doi: 10.1785/0220240309.

Supplemental Material

Introduction

Strong motion data recordings are essential for many different purposes in seismological and engineering applications. In particular, empirical parameters derived from observations are used to calibrate ground-motion prediction models, commonly referred to as GMMs, which are used for many applications, including the production of probabilistic (Probabilistic Seismic Hazard Assessment, PSHA) or deterministic (scenario analysis

1. Istituto Nazionale de Geofisica e Vulcanologia, Milan, Italy, https://orcid.org/0000-0001-7947-4281 (GL); https://orcid.org/0009-0003-3139-2887 (LV); https://orcid.org/0000-0003-2852-9144 (CF); https://orcid.org/0000-0002-5669-8650 (SS); https://orcid.org/0000-0003-4312-580X (LL); https://orcid.org/0000-0001-5745-0414 (FP); 2. German Research Centre for Geosciences, Potsdam, Germany, https://orcid.org/0000-0002-8619-2220 (DB)

*Corresponding author: giovanni.lanzano@ingv.it

or shaking maps) seismic hazard models. GMMs typically provide the median value of an intensity measurement and associated variability as a function of explanatory variables such as magnitude, distance, and site parameters. Where empirical data with good coverage of magnitude and distance are available, various techniques can be used to calibrate GMMs, adopting parametric (e.g., functional forms with and without random effects) and nonparametric (e.g., machine learning techniques) approaches. For this reason, many efforts have been made for more than 20 yr to collect accelerometer recordings of moderate-to-strong earthquakes, homogeneously processed and qualified with metadata of the highest possible quality.

Initially, parametric tables of waveform intensity measures and associated metadata were built on a continental or global scale, and only rarely on a national scale. These datasets are geared toward the calibration of large-scale GMMs, which are mainly based on the use of accelerometer recordings. At this stage, the Next Generation Attenuation (NGA)-West (v.1 Chiou *et al.*, 2008; v.2 Ancheta *et al.*, 2014) databases were built, with the aim of collecting strong motion recordings of world-scale events, to characterize the motion near the source, and weak motion recordings to describe the regional characteristics of seismic motion in the best monitored areas (mainly California, but also Italy, Japan, Türkiye, and others).

In the European context, the dataset of observations was progressively increased, starting with ISESD (Ambraseys *et al.*, 2004), continuing with RESORCE (Akkar *et al.*, 2014), up to the flatfile derived from European Engineering Strong Motion (ESM) (Bindi *et al.*, 2019; Lanzano, Sgobba, *et al.*, 2019). These databases have been the basis for different generations of GMMs for their use in the framework of European-scale hazard assessments (from Ambraseys *et al.*, 2005 to Kotha *et al.*, 2020).

More recently, the increased availability of records, thanks to the strong growth of monitoring networks, led to the construction of regional-scale datasets, also introducing velocimeter records and lowering the minimum magnitude (generally down to 3.0). These efforts have been made, considering different spatial scales and magnitude ranges, for Japan (Dawood *et al.*, 2016; Bahrampouri *et al.*, 2021), Chile (Bastías and Montalva, 2016), Türkiye (Zare and Bard, 2002; Akkar *et al.*, 2010), and more recently for France (Traversa *et al.*, 2020), Greece (Margaris *et al.*, 2021), and New Zealand (Hutchinson *et al.*, 2024).

A flatfile at the Italian scale has not yet been officially presented, partly because data from Italian earthquakes have been included in both the world (NGA-West) and European-scale databases. In fact, the specific characteristics of the seismicity on the Italian territory (mainly consisting of normal faults) and the high quality of the monitoring systems in Italy made the inclusion of these data crucial.

This article presents the main characteristics and preliminary analyses of a high-quality strong-to-weak motion flatfile specific for Italy, with the aim of providing a tool to perform seismological and engineering analyses. The dataset has been named ITACAext because it is intended to be an extended version of ITalian ACcelerometric Archive (ITACA; Felicetta *et al.*, 2023) database, including also a significant number of velocimetric recordings and waveforms from networks of countries bordering Italy.

A preliminary version of ITACAext (v. 1.0; Brunelli *et al.*, 2022) was already published on the ITACA website in 2022, including waveforms recorded up to the end of 2020. Based on this first release, a residual analysis has already been performed to investigate data quality, to estimate empirical amplification functions at the recording sites (ITACAs2s; Lanzano *et al.*, 2024) and to define a data-driven zonation for ground-motion modeling purposes (Brunelli *et al.*, 2023).

Data Collection

The ITACAext dataset is built by collecting the recordings available from some authoritative sources:

1. the accelerometric waveforms of the latest release of ITACA database (v 4.0; Felicetta *et al.*, 2023), which includes data from **M** > 3.0 events that occurred in Italy and in surrounding countries from 1972 to 2022;
2. the velocimetry records of short-period (EH), broadband (HH), and down-hole (DH) instruments from Italian monitoring networks (permanent and temporary), available on European Integrated Data Archive (EIDA) (Strollo *et al.*, 2021), for 666 earthquakes (9285 HH; 3360 EH; and 90 DH) (Fig. 1a); and
3. the accelerometric and velocimetric records of non-Italian seismic networks of neighboring countries (France, Switzerland, Slovenia, Greece, Albania, Croatia, and Malta) to complete the azimuthal distribution of events that occurred close to the border. In particular, the accelerometric recordings are collected from the ESM database (Luzi *et al.*, 2020), whereas the velocimetric data are collected from EIDA.

Note that a check against saturation of velocimetric records is carried out, adopting a conservative threshold to the number of counts for the waveforms downloaded from EIDA: if even one counts number exceeds 5,500,000, then the waveform is discarded.

The recording networks that contribute the most to ITACAext are (Fig. 1b):

- Italian National Seismic Network (INSN), code IV (Istituto Nazionale di Geofisica e Vulcanologia [INGV], 2005), which includes all the permanent and several temporary stations (both accelerometers and velocimeters), managed by the INGV (23,664; 45%);

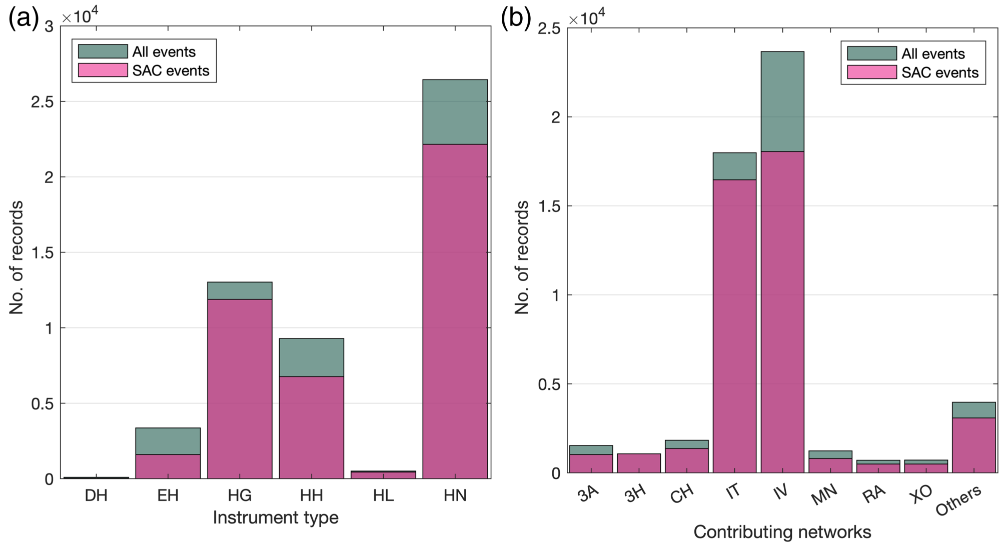


Figure 1. ITalian ACcelerometric Archive (ITACAext) data distribution according to: (a) instrument type; (b) monitoring network. SAC, shallow active crust. The color version of this figure is available only in the electronic edition.

- Rete Accelerometrica Nazionale (RAN), code IT (DPC, 1972) managed by the Italian Department of Civil Protection (17,981; 34%);
- Swiss Seismological Network (code CH; Swiss Seismological Service [SED] At ETH Zurich, 1983), comprising all permanent high-gain (short-period and broadband) and low-gain strong motion stations in Switzerland (1,829; 3.5%);
- temporary network of the Italian Seismic Microzonation Centre (code 3A, Cara *et al.*, 2019), installed in 2016 to monitor the seismic effects during the central Italy seismic sequence, managed by a consortium of different institutions, under the coordination of the Italian Seismic Microzonation Centre (1,537; 2.9%);
- Mediterranean project Network, MEDNET (code MN; MedNet Project Partner Institutions, 1990), composed of broadband stations in the Mediterranean countries and also managed by INGV (1,239; 2.4%);
- temporary network in the Norcia basin (code 3H; Pilz and Parolai, 2009), managed by the German Research Centre for Geosciences (GFZ), which includes several records of 2009 L'Aquila sequence (1,075; 2.0%);
- temporary seismic network for site effect studies in the Amatrice area (central Italy) (SESAA) (code XO, EMERSITO Working Group, 2018), installed shortly after the 2016 Amatrice mainshock by the INGV Emersito Working group (714; 1.4%); and
- Réseau Accélérométrique Permanent (French Accelerometric Network) (code RA; RESIF, 1995), managed by RESIF (Reseau sismologique et géodésique français) (704; 1.3%).

The full list of networks can be found in Appendix A.

After the collection, the dataset consists of 52,706 recordings from 2,338 events and 2,294 stations. Figure 2a shows the spatial distribution of the events in the dataset, which are mainly distributed along the Apennine arc, in eastern Sicily, in the southern Tyrrhenian Sea and, to a lesser extent, in parts of the Po Valley, the Adriatic Sea, and the Alps. On the other hand, the available stations are fairly evenly distributed throughout the country (Fig. 2b), with the exception of Sardinia. The density of stations is very high in the central Apennines and less so in parts of Apulia, Tuscany, and northern Italy (Piedmont, Lombardy, and Veneto).

Figure 2c shows the distribution of recordings in ITACAext over time, both in terms of cumulative number of records and by calculating the average number of recordings for events in each year. The major contribution, with more than 10,000 recordings, comes from the records of events that occurred in 2016, corresponding to the central Italy seismic sequence, but a nonnegligible contribution comes from other sequences such as the Po Plain 2012 and L'Aquila 2009. The number of records per event increased progressively, starting from very few registrations per event (2–5), up to 40–80 registrations after 2016; the increase accelerated sharply from 2008 also thanks to the expansion of the monitoring networks in Italy.

Intensity Measures

The waveforms in the flatfile were uniformly processed following the procedure described in Paolucci *et al.* (2011) and detailed in Pacor *et al.* (2011), which prescribes the application of a second-order acausal time-domain Butterworth filter to the zero-padded acceleration time series and zero-pad removal to make acceleration and displacement consistent after double integration. A low-pass filter with a fixed frequency of 30 Hz was applied to the three components of the ground motion; the high-pass frequencies depend on magnitude ranges (table 1 of Puglia *et al.*, 2018). Then, the waveforms were reviewed by experts, and the default values were modified. Figure A1 (available in the supplemental material) shows the final values of the ITACAext filtering frequencies compared with the default values, showing that can be significantly different. The manual revision also allowed to identify and exclude poor-quality records, that is, noisy records or time histories containing spikes (Vitrano *et al.*, 2024). To keep track of human intervention, the band-pass corner frequencies associated with each record component are given in the table.

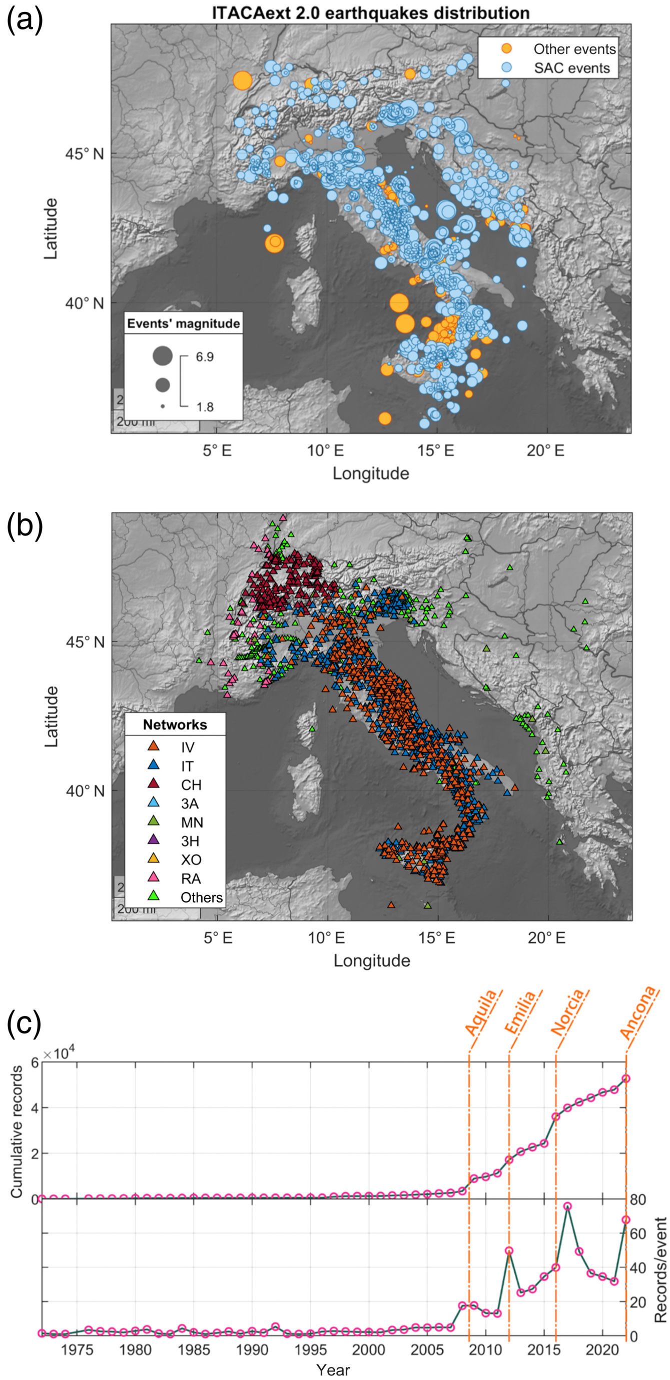


Figure 2. Spatial and temporal distribution of ITACAext. (a) Map of earthquakes with the symbol size scaled by moment magnitude; (b) map of the recording stations colored by network; and (c) number of recordings as a function of time. Top: cumulative number of recordings per year; bottom: average number of recordings per event per year. The color version of this figure is available only in the electronic edition.

The flatfile table contains the peak ground-motion measures of acceleration (PGA), velocity (PGV), and displacement (PGD) time histories, some integral and duration parameters, that is, the Housner intensity, the cumulative absolute velocity (CAV), the Arias intensity (AI), and the significant duration (Trifunac and Brady, 1975). The 5% damping elastic response spectral ordinates in acceleration (SA) are also provided at 36 periods ranging from 0.01 to 10 s. These parameters are provided for each record component (two horizontal and one vertical) and in terms of orientation independent measures, RotDnn (Boore, 2010), that is, the RotD50 (the median), the RotD100 (the maximum), and the RotD00 (the minimum) of the intensity measure distribution over all rotation angles.

In addition, the ordinates of the Fourier amplitude spectrum (FAS) in acceleration in the 0.04–50 Hz frequency range are also provided, using a Konno and Ohmachi (1998) function with a bandwidth smoothing coefficient of $b = 40$. The aim is to provide data useful both for the calibration of empirical prediction models in FAS and to constitute a reference dataset for performing generalized inversion analyses in Italy. For this reason, parametric tables of Fourier spectrum amplitudes are provided both for the whole signal and for a portion of the signal containing only the "strong" phase, based on *S*-wave arrivals.

The signal cut is based on an energy criterion, that is, on the identification of time instants corresponding to different percentages of the total AI. The first cut time T1 is obtained according to the following criteria:

- if the signal is late triggered (for older analog signals triggered on *S* waves), 1% of the total AI is used;
- if the station has an epicentral distance < 10 km, 2.5% AI is used; and
- in the remaining cases, 5% AI is assumed.

For the sake of simplicity, the final T1 is unique for the three components and is given by the average of the T1-values calculated for the two horizontal components of the signal.

The second cut time T2 is estimated as follows:

- if the station has an $R_{epi} < 25$ km, 95% AI is used;
- if the station has a $25 \text{ km} < R_{epi} < 50$ km, 90% AI is used; and
- in the remaining cases, 85% AI is used.

The final value of T2 is also obtained by averaging the T2 calculated for the two horizontal components of the signal. Finally, it is checked whether the total duration of the cut, that is, $dT = \text{T2}-\text{T1}$, is less than 5s; in these cases, a fixed $dT = 5$ s is conservatively assumed and the value of T2 is adjusted accordingly. Two examples of the cuts that are applied to the signal, including the median prediction obtained from the Bommer *et al.* (2009) model for the significant duration, are reported in Figure A2. After the cut, a cosine taper is applied to the waveform and spectral amplitudes are computed for about 100 frequencies, logarithmically sampled in the interval 0.04–50 Hz. The table provides amplitude values for each component (two horizontal and one vertical).

Because the processing is manual, the corner frequencies of the high- and low-pass filters are selected by experts for each individual record (the values are also provided in the flatfile). In fact, the number of usable records is different for each spectral ordinate. In the interval 0.7–10 Hz, all spectral ordinates in FAS and SA can be used, whereas for frequencies below 0.7 Hz, the number gradually decreases, reaching 60% of the total at 0.2 Hz and 19% at 0.1 Hz. At high frequencies ($f > 10$ Hz), a similar decrease is observed for the number of FAS ordinates: 48% of the data can be used at 30 Hz and only 1% at 50 Hz. On the other hand, the SA short-period ordinates are insensitive to the choice of corner filter (Douglas and Boore, 2011) and can be considered without curtailment in ground-motion analyses.

Metadata Attribution for Seismological Analyses

The flatfile includes many fields related to event, station, and waveform metadata, largely replicating the structure of the ESM flatfile (Lanzano, Sgobba, *et al.*, 2019). For convenience, we provided, as the supplemental appendix, ESUPP2, the user manual of the flatfile, also available via the ITACA website, in which the meanings of all table fields (data and metadata) are explained. Metadata in ITACAext are carefully reviewed and supporting information is provided, such as an appropriate reference or the type of analysis/test used to obtain such parameters. However, in some cases, such metadata may not be homogeneous because they are derived from authoritative sources that may provide incomplete information.

In this section, we illustrate some criteria that were used to qualify all the waveforms in terms of the main parameters used as explanatory variables in ground-motion analysis, namely magnitude, distance, focal mechanism, focal depth, and site characterization parameters.

The table contains several fields with different magnitude estimates from those available in the database, including moment M_w, local M_L, surface waves M_s, and body waves M_b, duration M_d, or a generic magnitude **M** if the authoritative source does not specify the origin of the magnitude (Fig. 3a). Because prediction models are generally calibrated to moment magnitude, which is not affected by saturation for strong earthquakes such as M_L, other magnitude estimates have been normalized to this reference. Furthermore, Gasperini *et al.* (2012) has pointed out that there is also a scaling between moment magnitude estimates obtained from different methods in Italy, namely the Regional Centroid Moment Tensor (RCMT; Pondrelli *et al.*, 2002) and the Time-Domain Moment Tensor (TDMT; Dreger *et al.*, 2003). We refer to the RCMT as the target M_w estimate, and use the empirical relation proposed by Brunelli *et al.* (2023), valid for shallow active crustal (SAC) events with **M** > 3.5, to convert M_w–TDMT into M_w–RCMT. We adopt this reference estimate to make our analysis consistent with the approach used to calibrate the GMM for Europe (Kotha *et al.*, 2020) by adopting the European-Mediterranean earthquake catalogue (EMEC) magnitude (Grünthal and Wahlström, 2012; Lammers *et al.*, 2023). All the EMEC magnitudes were converted to M_w using regional relations.

For events where no M_w estimates are available, the Gasperini *et al.* (2013) relationship is used to convert the M_L into M_w–RCMT. For events where neither M_w nor M_L estimates are available (<1%), the following hierarchy of choices is used: M_s, M_b, M_d, and generic **M**. Because there are sometimes no conversion equations between these metrics and M_w–RCMT, we decided to consider them without multiple conversions to avoid introducing further uncertainty. In the parametric table, we provided this standardized M_w–RCMT estimate for all records flagged as SAC, according to the selection procedure described in the Selection for SAC section. The histogram of the M_w–RCMT vector is shown in Figure 3b: as expected, the majority of the data has a magnitude between 3.5 and 4.5 (61%); the contribution of events with M_w–RCMT < 3.5 is lower because the sampling of these events is not complete (16%); the percentage of events with M_w–RCMT > 4.5 is drastically reduced (23%).

The focal depth of the events has a large influence on the seismic amplitudes but is subjected to considerable uncertainty. A significant improvement was achieved in the construction of ITACAext, thanks to the introduction of the revised localizations made available in the CAT3 catalog (Spallarossa *et al.*, 2021) for the events in central Italy. Figure 3c shows the distribution of data with respect to the type of catalog used to locate them, dividing them into three broad categories: global catalogs (European-Mediterranean Seismological Centre, EMSC; International Seismological Centre [ISC], 2020), regional catalogs (INGV bulletin, ISIDe, 2007), and local catalogs (CAT3 or local catalogs of volcanic events). By updating the event metadata in ITACA, about 30% of the data could be qualified with a depth estimate obtained from the best-quality locations of local catalogs. However, the depth distribution still has a sharp peak at 10 km, corresponding to the conventional depth assigned by near-real-time automatic systems of regional or global localization (Fig. 3d).

For most moderate-to-strong earthquakes, the focal mechanism is primarily obtained from the moment tensor solution, which is consistent with M_w estimates. Where these are not available, especially for the oldest and low-magnitude events, the style of faulting is inferred based on those proposed by available seismogenic zonations. In this study, we refer to ZS16 (Visini *et al.*, 2022), which provides a most likely style of faulting for each seismogenic zone, based on the study by Pondrelli *et al.* (2020). Figure 3e shows the distribution of the data based on the method by which the fault mechanism was associated with the event: only 2% of the events, almost all with **M** > 4.5, have an estimate of the focal mechanism from the moment tensor solution. Figure 3f, on the other hand, confirms that most of the events in the dataset have a normal focal mechanism because the dataset is dominated by events in central Italy.

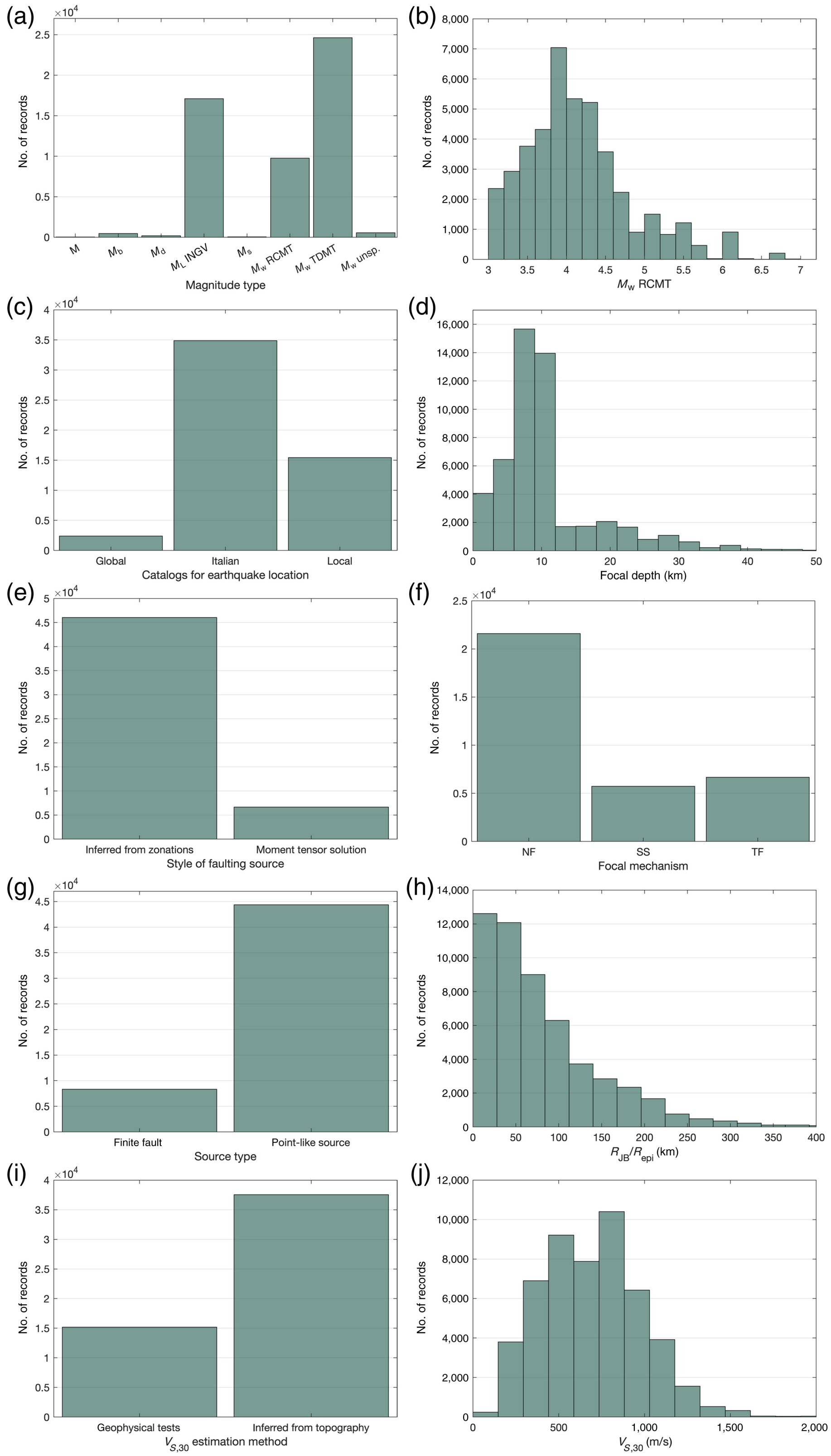


Figure 3. ITACAext data distribution for different explanatory variables. (a) magnitude type; (b) RCMT moment magnitude estimates; (c) catalogs for earthquake location; (d) focal depth estimates; (e) sources for style of faulting; (f) style of faulting estimates; (g) sources for finite faults; (h) Joyner-Boore or epicentral distances values; (i) methods for the estimation of $V_{S,30}$; (k) $V_{S,30}$ values. The color version of this figure is available only in the electronic edition.

Ground-motion prediction models require distance metrics that take fault geometry into account, such as the Joyner–Boore distance, R_{JB}, and the distance to the rupture plane, R_{rup}. ITACAext provides the fault geometries for all events with magnitude greater than 5.0 (Fig. 3g); in the remaining cases, that is, for small earthquakes, the size of the faults is such that the differences between R_{JB} and R_{epi} can be assumed to be minimal, such as those between R_{rup} and R_{hypo}. Figure 3h shows that a significant amount of the records in the dataset are within 20 km (16%), making the dataset useful for improving the description of near-source motion characteristics. In particular, the range of horizontal PGAs of “near-source” records is pretty wide (10^{-1}–10^3 cm/s^2), and the spatial distribution of stations with these records covers large part of the Italian territory (see Fig. A3).

The most important parameter for site characterization in the GMMs is the $V_{S,30}$, that is, the average shear-wave velocity in the top 30 m (European Committee for Standardization [CEN], 2003). $V_{S,30}$-values are generally calculated from V_S profiles with depth, available at least up to 30 m above the ground level, obtained from borehole or surface geophysical tests. ITACAext benefits from the collection of V_S profiles in the ITACA and ESM databases, retrieved from various projects and appropriately referenced in the flatfile. However, these types of surveys are not available for all sites (about 30% according to Fig. 3i), so it is necessary to use proxies to infer the value of $V_{S,30}$. In the flatfile, estimates

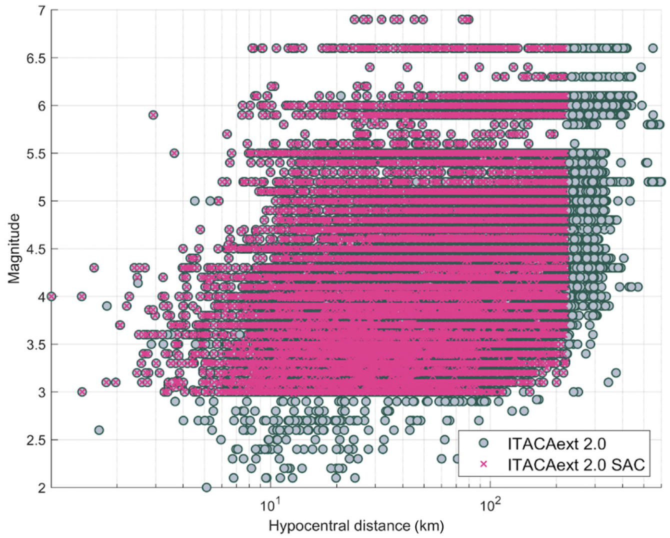


Figure 4. Magnitude–distance scatter plot. The color version of this figure is available only in the electronic edition.

of $V_{S,30}$ and/or EC8 classes from two different proxies are available: (1) correlation with surface geology, consisting of the classification of sites into EC8 classes, on the basis of the lithological classification of surface material; (2) correlation between $V_{S,30}$ and slope, first proposed by Wald and Allen (2007) and adapted to Italy by Mascandola *et al.* (2021). In the case of geologic proxy, we can assign a representative $V_{S,30}$ for each EC8 class, on the basis of the average of the measured $V_{S,30}$ for each class ($V_{S,30} = 1000$ m/s for EC8-A, 550 m/s for EC8-B, 270 m/s for EC8-C, and 150 m/s for EC8-D). The estimates of $V_{S,30}$ from topographic slope, on the contrary, have the advantage of varying with continuity in space, suggesting a priority over geology in the hierarchy of assigning $V_{S,30}$ to ITACAext sites. For very few stations where the slope values are missing a reference $V_{S,30} = 550$ m/s can be assigned, corresponding to the central value of the $V_{S,30}$ interval of EC8-B soil category, which is the most represented category among the sites with measured $V_{S,30}$ (Fig. 3j).

Figure 4 shows the magnitude (M_w–RCMT)–hypocentral distance distribution of the dataset: the data are well-sampled in the magnitude range 3.0–6.1 and for epicentral distances from 20 to 250 km. There is also a significant number of records related to strong events with magnitudes between 6.0 and 6.9, corresponding to about 10% of the records.

Selection for SAC Regions

Preliminary selection

Earthquakes in Italy can have different characteristics because they can be associated with different types of seismicity, such as active crustal seismicity, volcanic activity, subduction areas, deep seismicity, and even stable continental activity (Sardinia). However, the predominant and most frequent seismicity on Italian territory is related to active faults that are distributed over almost the entire national territory and are responsible for the most destructive earthquakes of the last 100 yr. For this reason, we include a flag ("SAC_events") in the table to identify records related to active crustal events. The criteria used for this selection are strongly oriented toward seismological and engineering applications such as the calibration of seismic motion prediction models or performing the spectral decomposition of the ground motion (e.g., Generalized Inversion Technique or simply GIT). A summary of all the steps that led to the final selection is given in Table 1.

The starting selection criteria are:

- moment magnitude $M_w > 3.0$;
- focal depth < 35 km;
- epicentral distance up to 220 km; and
- surface sensors or down to 3 m depth.

The criterion based on the focal depth allows to exclude earthquakes in the subduction zones of the lower Tyrrhenian Sea, which are characterized by a much deeper seismicity than the crustal zones. On the other hand, volcanic events, which occur at the same depths as crustal earthquakes, must be excluded using a different strategy. The source zonation by

TABLE 1
Steps for the Construction of the Shallow Active Crustal (SAC) Subset

Collection Steps	Description	Records Number
#1	Initial collection (see the Data Collection section)	52,706 (100%)
#2	Identification of shallow crustal events	45,714 (87%)
	Basic selection compatible with the GMMs application validity	
	Removing problematic waveforms from previous studies (see the Preliminary selection section)	
#3	Selection after the colocated sensors comparison (see the Comparison of colocated stations amplitudes section)	42,887 (81%)
#4	Consistency check against an ad hoc GMM (see the Consistency check and final selection section)	42,877 (81%)

Brunelli *et al.* (2023) is considered, which includes a number of zones classified as volcanic, corresponding to the districts of the Colli Albani, Vesuvius, Campi Flegrei, Ischia, Etna, and the Aeolian islands. For the sake of simplicity, events falling within these zones are simply removed from the selection.

In addition, some problematic waveforms have been excluded, based on the results of previous studies (the complete list is provided in the supplemental appendices ESUPP3):

- the outliers provided by the Mascandola *et al.* (2022) study, analyzing the ITACAext v.1.0 dataset (Brunelli *et al.*, 2022);
- the recordings of the 9 November 2022 M_w 5.2 earthquake with the ITACA code “INT-33301931,” in the tail of the mainshock (M_w 5.5) of the Ancona sequence of November 2022;
- the recordings of station SN1 of the IV network, installed on the seafloor of the Ionian Sea; and
- the recordings of some stations that had ambiguous location codes or are not related to free-field stations, that is, all recordings of stations NCAD (RA network) and SBUS (CH network), all stations with instrument code DH and location code 01 (located in borehole), and the recordings of station RIBO (IV network) with location code 01.

Comparison of colocated stations amplitudes

After applying standard filters to identify crustal events, a quality/consistency analysis of the dataset was performed based on a two-step residual analysis. The total residuals are calculated as

$$R_{es} = \log_{10} Y_{\text{obs}} - \log_{10} Y_{\text{pred}}, \quad (1)$$

in which Y_{obs} are the seismic motion parameters obtained from the ITACAext observations after the preliminary selection and Y_{pred} are the predictions of a reference GMM model.

In the first step, the residual analysis is carried out to support the choice between colocated instrument recordings. Indeed, it is not only very common in the IV network but also in other permanent and temporary networks (3A, MN, IX, among the others) to arrange measuring stations with colocated instruments, usually an accelerometer (usually HN) and a velocimeter (often HH, sometimes EH). With regard to the use of ITACAext data for applications such as the calibration of ground-motion prediction models, the presence of potentially duplicate data could overweight these observations in the mathematical algorithms used. For this reason, we defined a selection procedure supported by the results of the residual analysis.

For this first stage analysis, the parameters for which residuals are calculated are the horizontal ordinates of the Fourier spectrum in the range 0.1–30 Hz. The reference GMM is the one proposed by Lanzano *et al.* (2022) for Italy (see Appendix II for details on GMM). The total residuals R_{es}, according to equation (1), are then decomposed, applying mixed-effects regression (Bates *et al.*, 2015), according to the following scheme:

$$R_{es} = \delta_0 + \delta B_e|\text{event} + \delta I2I_s|\text{sensor} + \delta W_{es}, \quad (2)$$

in which δ_0 is the total residuals offset; δB_e is the between-event term, the average deviation away from the median prediction of the GMM for any event e; $\delta I2I_s$ is an instrument-to-instrument term, which defines the systematic bias of ground motions recorded by each sensor s with respect to the median GMM prediction; δW_{es} is the event- and instrument-corrected residual, that is, the remaining aleatory residual, from which the repeatable effects of event and instrument have been removed.

In this application, the classical decomposition scheme (event, site, and aleatory residuals as in Al Atik *et al.*, 2010 and many other studies) was modified by replacing the station term with a sensor systematic term. If the recording station is equipped with only one sensor, then the repeatable term of the sensor corresponds to the station term; if, however, the station is equipped, or has been equipped, with more than one sensor, then a specific contribution can be derived for each sensor. The advantage of applying this approach is that it enables the site effect, which has been shown to be represented by the repeatable station term, to be estimated separately for each sensor, even if the colocated sensors were operated at two different, not necessarily overlapping, epochs. The rationale behind this approach draws also inspiration from the study by Hollender *et al.* (2020), which shows that instruments close located with different installation types (as is commonly done for accelerometers and velocimeters) may show different site responses, especially at high frequencies. In addition, Cauzzi *et al.* (2022) took advantage of the large availability of recordings from colocated strong motion and broadband sensors in EIDA to develop strategies for detecting anomalies in both waveform data and station metadata.

Figure 5a,b shows two examples of $\delta I2I_s$ for colocated stations (IV.BDI and IV.CERA, in which the name is a combination of network and station codes), which were calculated for convenience with respect to a fixed reference motion (equal to $V_{S,30} = 800$ m/s, named $\delta I2I_{800}$) to make them analogous to an amplification function with respect to an engineering bedrock (Lanzano *et al.*, 2022). The two selected stations, both equipped with an HN accelerometer sensor and an HH broadband sensor, show very similar $\delta I2I_{800}$ trend from 0.1 to about 10 Hz, while slightly differ for higher frequencies. The maximum difference ($f \sim 25$ Hz) is in absolute value of the order of 0.4 ln units, corresponding to a reduction in amplitudes of about 30%.

To provide more robust statistics of the observed cases, the deviation Δ between the $\delta I2I_{800}$ of the accelerometer and velocimeter, frequency by frequency, is calculated as

$$\Delta(f) = \delta I2I_{800}^{\text{acc}}(f) - \delta I2I_{800}^{\text{vel}}(f). \quad (3)$$

The median and standard deviation of the differences Δ of the colocated accelerometer–velocimeter stations are then

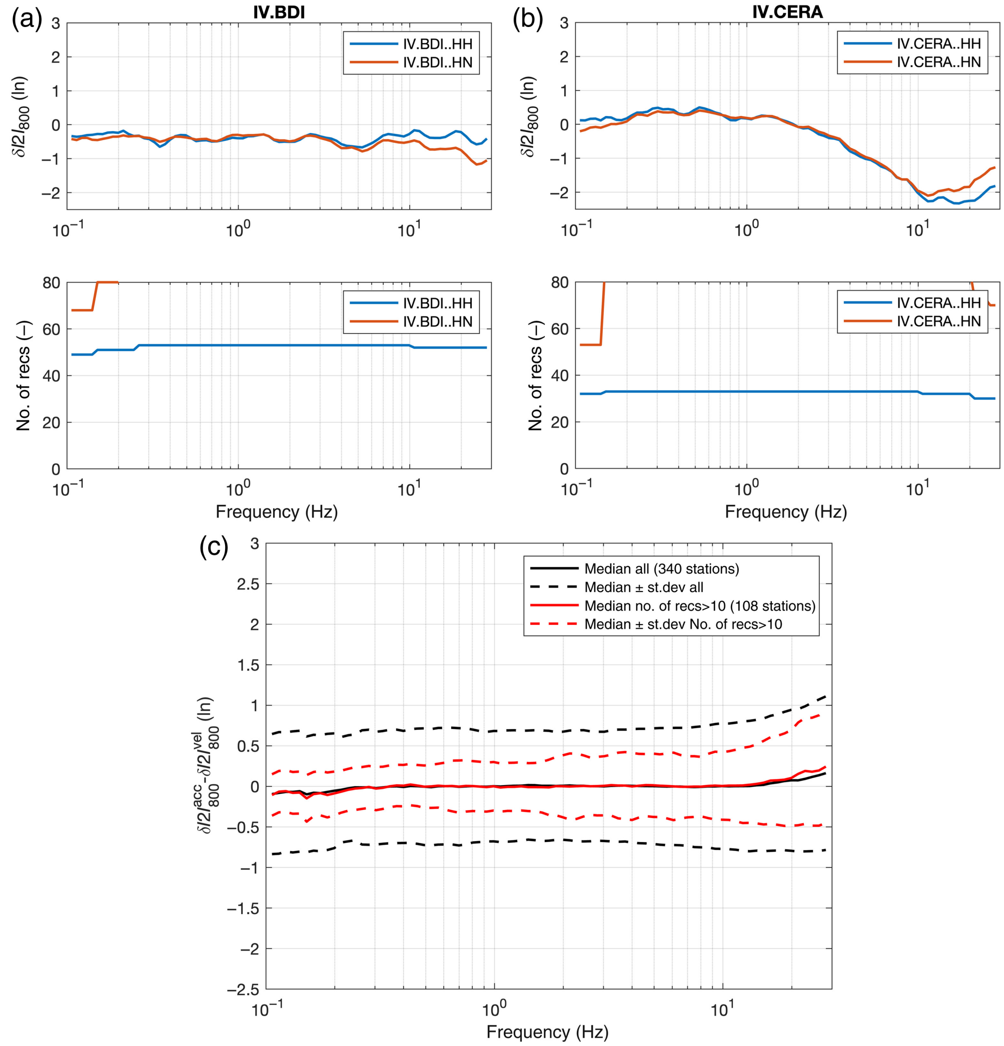


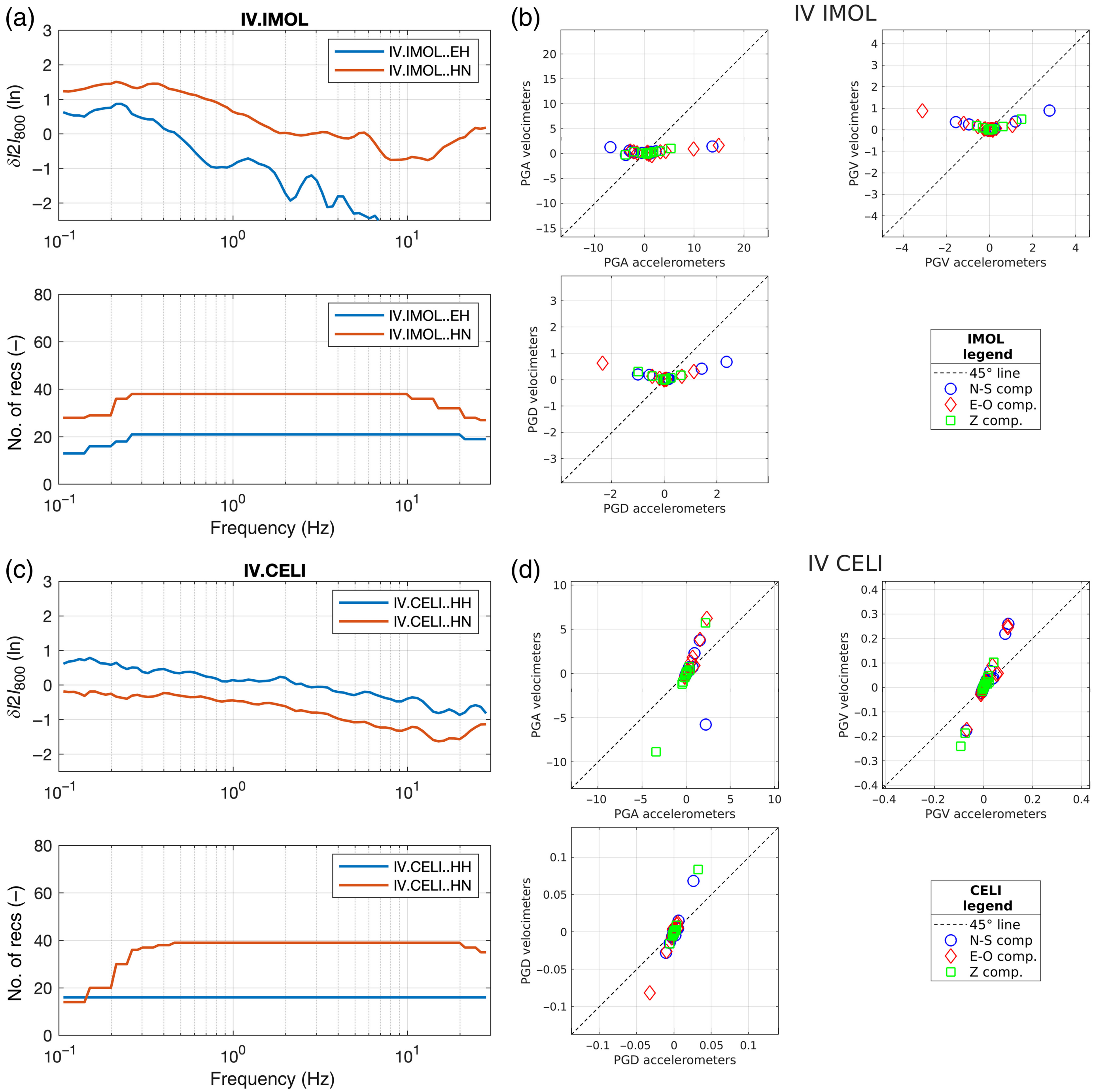
Figure 5. Fourier amplitude spectrum (FAS) instrument-to-instrument residual term (δI2I800) and number of records as a function of frequency. (a) IV.BDI; (b) IV.CERA. (c) Median value of Δ (difference between colocated accelerometer and velocimeter δI2I800 according to equation 3) as a function of frequency. The color version of this figure is available only in the electronic edition.

calculated, taking into account all data and cases in which both colocated sensors recorded more than 10 events (not simultaneously). The results, shown in Figure 5c, confirm that the median responses of the sensors are consistent over a wide range of frequencies from 0.25 to 15 Hz. Moreover, when each sensor is sampled from a sufficient number of recordings, the associated variability is strongly reduced, especially at low frequencies.

Systematic differences are observed at frequencies $f < 0.25$ Hz, with weak negative Δ, and at $f > 15$ Hz, with positive Δ to a median reduction in the amplitude of velocimetric ground motion of about 20% at 25 Hz. Moreover, the increase in variability at high frequencies indicates that the differences between the colocated sensors become progressively larger as f increases.

Focusing on the outliers, there are several cases (65 colocated stations), manually revised, in which the $\delta I2I_{800}$ trends remarkably diverge in different frequency ranges. The casuistry is quite wide, but we group the various cases into a few types:

- Some stations show $\delta I2I_{800}$ with significantly different responses between accelerometer and velocimeter, in which either curve shows a strongly amplified or deamplified systematic sensor response. As an example, Figure 6a shows the case of the IV.IMOL station, for which the short-period $\delta I2I_{800}$ exhibits considerable deamplification at high frequencies (Fig. 6b). These stations are manually examined, submitted to the check of network operators, and the sensor with a suspected nonphysical response is excluded. In the case of IV.IMOL, the velocimeter sensor (EH) was from an instrument installed in the borehole, and the information was not correctly reported in the station metadata.
- Some other stations exhibit a constant shift between the $\delta I2I_{800}$ curves of accelerometer and velocimeter, as shown in the example in Figure 6c related to station IV.CERA. In this case, the amplitudes of the accelerometer are proportionally smaller than those of the velocimeter (Fig. 6d). This behavior suggests possible differences and possible errors in the scale factors to physical units for the two sensors, which are also reported to the network operators. In cases in which there is no indication for the choice of sensor, we calculate the median value of the total residuals of the two sensors' recordings for the FAS amplitudes. The sensor selected is the one with the lower bias averaged over all frequencies.
- Some recording stations of the 3A network, such as station 3A.MZ12 in Figure 7, show systematic differences for $\delta I2I_{800}$-values at very low frequencies (<0.2 Hz). In particular, EH sensors show significant amplification compared to HN-values, which does not seem to be caused by any



site/installation effects. In these cases, we assume that the amplitudes of EH at low frequencies should be excluded in favor of those of HN.

The complete list of the excluded sensors can also be found in the supplemental appendix. For the rest of the colocated instruments, the following rules are adopted:

- if HN/HG and HH registrations are available for the same event, then HH is preferred;
- if HN/HG and EH recordings are available for the same event, then HN/HG is preferred;
- for EH recordings of the 3A network, the low cut was set to 0.2 Hz, if lower; and
- for some stations of the IX network (Irpinia Seismic Network ISNet, managed by University of Naples) equipped with two

Figure 6. Examples of inconsistent stations between colocated instruments: (a,b) IV.IMOL; (c,d) IV.CELI. (a,c) FAS instrument-to-instrument residual term (δl2l800) and number of records as a function of frequency; (b,d) peak parameters comparison (peak ground acceleration [PGA], peak ground velocity [PGV], and peak ground displacement [PGD]) of records of the same event from two colocated instruments. The color version of this figure is available only in the electronic edition.

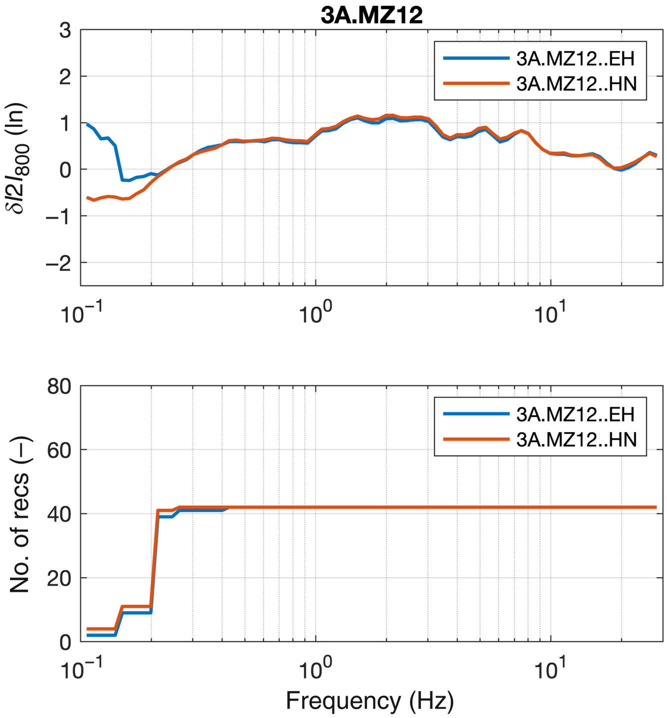


Figure 7. FAS Instrument-to-instrument residual term (δI2I800) and number of records as a function of frequency for the station 3A.MZ12. The color version of this figure is available only in the electronic edition.

accelerometers HN, we select the recordings with location code "'01."

Consistency check and final selection

In the second step, a basic consistency check is carried out for the identification of outliers in the dataset. This procedure is quite common in the preparation of datasets for seismological analyses (Bindi *et al.*, 2019; Traversa *et al.*, 2020; Mascandola *et al.*, 2022) and aims to recognize events, stations or recordings with anomalous behavior, which cannot be attributed to peculiar aspects of seismic motion (systematic amplification or deamplification).

The residuals are the result of calibrating a "base" median model (with the same functional form as ITA18) on the dataset in PGA, PGV, and two spectral ordinates (the same as Bindi *et al.*, 2019). In this second analysis, an ad hoc model is calibrated with the aim of ensuring that the overall residuals have no median bias (offset in equation 3 $\delta_0 = 0$), allowing better identification of outliers on the tails of the residual distribution. For this reason, the coefficients and associated variabilities of the model are not provided. Information on the functional and calibration methods can be found in the article by Lanzano, Luzi, *et al.* (2019) and is briefly reported in Appendix I.

In this case, the total residuals are standardly decomposed into

$$R_{es} = \delta B_e|\text{event} + \delta S2S_s|\text{station} + \delta W_{es}, \quad (4)$$

in which δB_e has the same meaning as the term given in equation (4), whereas $\delta S2S_s$ is the site-to-site term, that is, the systematic bias of the ground motions recorded by each station s with respect to the median GMM prediction. Finally, δW_{es} is the event- and station-corrected aleatory residual.

Figure 8 shows the trends of the components of the residuals as a function of different explanatory variables for PGA and PGV: in particular, δB_e as a function of magnitude (M_w–RMCT), $\delta S2S_s$ as a function of $V_{S,30}$ (preferred estimate according to the hierarchy shown in the Metadata Attribution for Seismological Analyses section), and δW_{es} as a function of Joyner–Boore distance are shown. The corresponding plots for the spectral ordinates are available in Figure A4.

A preliminary check that there is no systematic bias in the distributions of the residuals with the explanatory variables has been made. However, it is clear that there are some empirical data (events, stations, or records) that can be considered as outliers. Reasonably, this occurs when δB_e and $\delta S2S_s$ exceed one decimal logarithmic unit in absolute value (one order of magnitude difference between observed and predicted) and about 1.5 $\log_{10}$ units for $|\delta W_{es}|$. Because the threshold for identifying outliers may be different for the parameters used as reference (in general, variability is higher for high-frequency parameters such as PGA), we decided to use a more general criterion. In particular, recurrent outliers (for at least three of the four intensity measures considered) with residuals several times greater than the standard deviation of the model calibrated to the data are excluded: records with residuals δW_{es} in absolute value exceeding 6 times the standard deviation ϕ_0 (standard deviation of δW_{es}); events with $|\delta B_e|$ exceeding 4 times the τ (standard deviation of δB_e); stations with $|\delta S2S_s|/\phi_{S2S} > 4$, in which ϕ_{S2S} is the standard deviation of $\delta S2S_s$. This second check leads to the exclusion of a further 50 records, which are also listed and flagged in ESUPP3.

As a result of the exclusion of these additional data, the final dataset (named ITACAext_sac) for the active crustal areas consists of 42,877 records, 2,052 events, and 2,072 stations. The overall reduction of the dataset compared to the initial collection is only 19%, confirming that a large part of ITACAext is composed of records of events in the active tectonic crustal environment, making it particularly suitable for the study of ground motions of this type. In addition, the magnitude–distance distribution is not particularly affected by ITACAext_sac selection (Fig. 5) because the largest magnitude earthquakes in the dataset and most of the near-source recordings are associated with events in the active crustal tectonic environment.

Figure 9a shows the spatial sampling of the source-to-site record paths in ITACAext_sac: the spatial coverage of the paths is very dense over much of the Italian territory and considerably exceeds that of the datasets used to calibrate regional (Lanzano, Luzi, *et al.*, 2019) and recent European (Kotha *et al.*, 2020) models. Some areas remain uncovered, such as parts of Apulia (Salento peninsula), Tuscany (coastal area), and

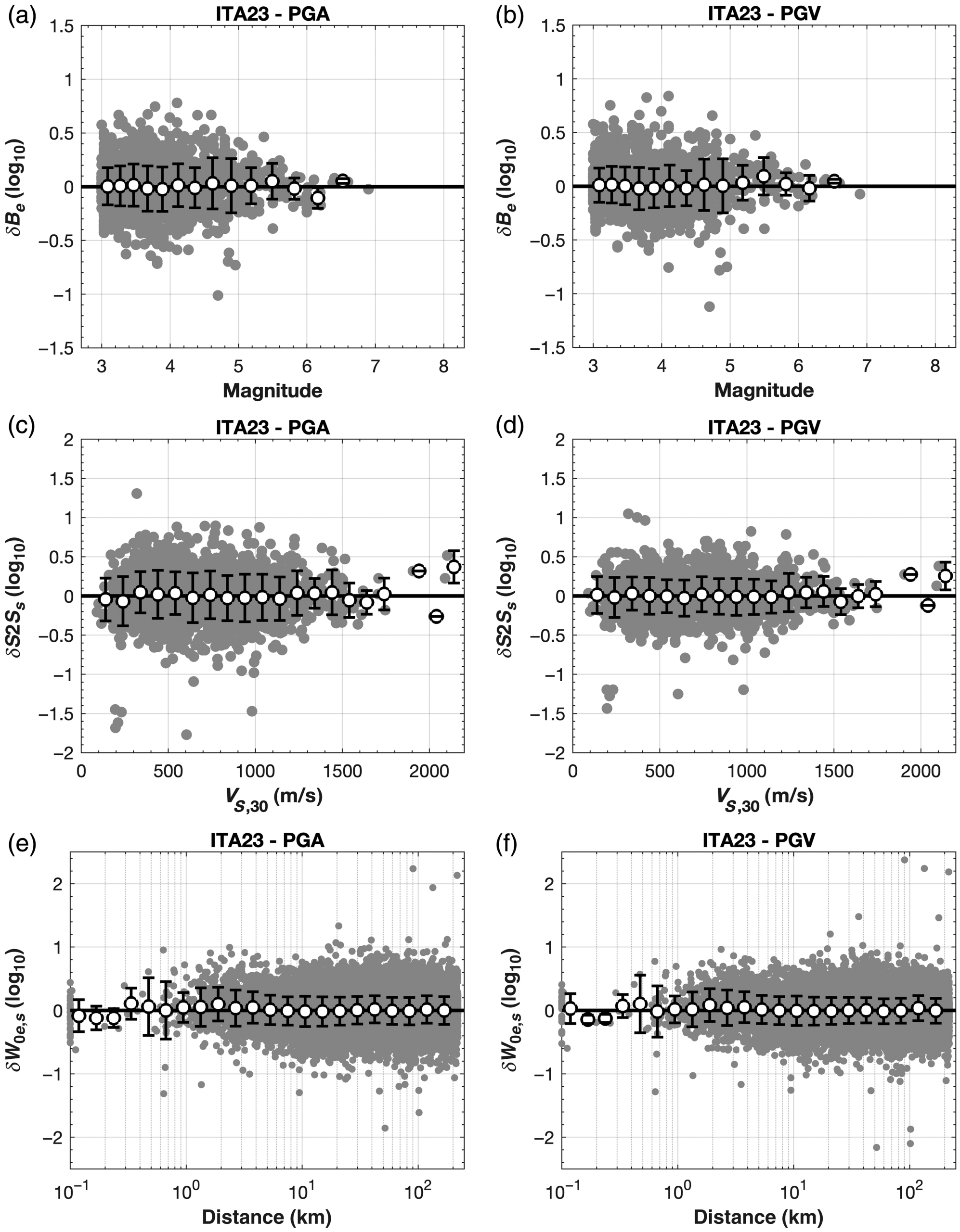


Figure 8. Residual decomposition of an ad hoc ground-motion model (GMM) for outlier detection (second-step residual analysis). (a,c,e) PGA; (b,d,f) PGV. (a,b) δB_e versus M_w–Regional Centroid Moment Tensor (RCMT); (c,d) $\delta S2Ss$ versus $V_{S,30}$; (e,f) δWes versus Joyner–Boore distance.

amplifications. Following the study of Bindi and Kotha (2020), we applied a nonparametric approach but, instead of considering a reference site amplification model, we use a reference event approach to break the trade-off between source and site contributions. In particular, the source spectral shapes of the 2009 L'Aquila and 2016 Norcia earthquakes are constrained to an omega-square model with source parameters obtained by Morasca *et al.* (2022). For further details on the procedure, see Bindi and Kotha (2020).

Figure 10 shows the obtained site amplifications, with values exceeding a factor of 10 for stations installed on sediment, and an almost flat trend suggesting the absence of site resonance effects for stations installed on rock. The results of a few well-studied cases are shown in color, such as the CLF station installed in the Colfiorito basin, which shows amplifications at 1 Hz due to basin-induced Love waves (Di Giulio *et al.*, 2003); the amplification at the NCR2 station in Nocera Umbra is related to a buried wedge of weathered rock (Rovelli *et al.*, 2002); the LSS (Leonessa Nova) and LLS (Linth–Limmern) stations installed on rock were considered as reference stations free from significant site effects by Morasca *et al.* (2023) and Bindi *et al.* (2023), respectively. The few stations with unrealistic amplifications (orange thin lines) are located in the Balkans, in which the ITACAext_sac data coverage is not suitable for the spectral decomposition (Fig. 9). The black thick line in Figure 10 corresponds to the CMB (Campobasso) station, which is installed at the base of a hill on a semiflat area. The GIT result shows strong amplifications at frequencies below 1 Hz, followed by deamplification at the higher frequencies, which could be due to topographic effects. Because no detailed geophysical survey is available for the

Lombardy (Lake Maggiore area), as well as the entire territory of Sardinia. Path density is also estimated by considering the coordinates of the midpoint between the source and site coordinates and counting the number of records for each cell of a 0.2° × 0.2° equispaced grid. The results show that more than 90% of the Italian territory has more than 10 paths with the midpoint falling within each cell.

Example of a GIT analysis

As a final analysis of the ITACAext_sac, a spectral decomposition GIT is performed to check the quality of the site

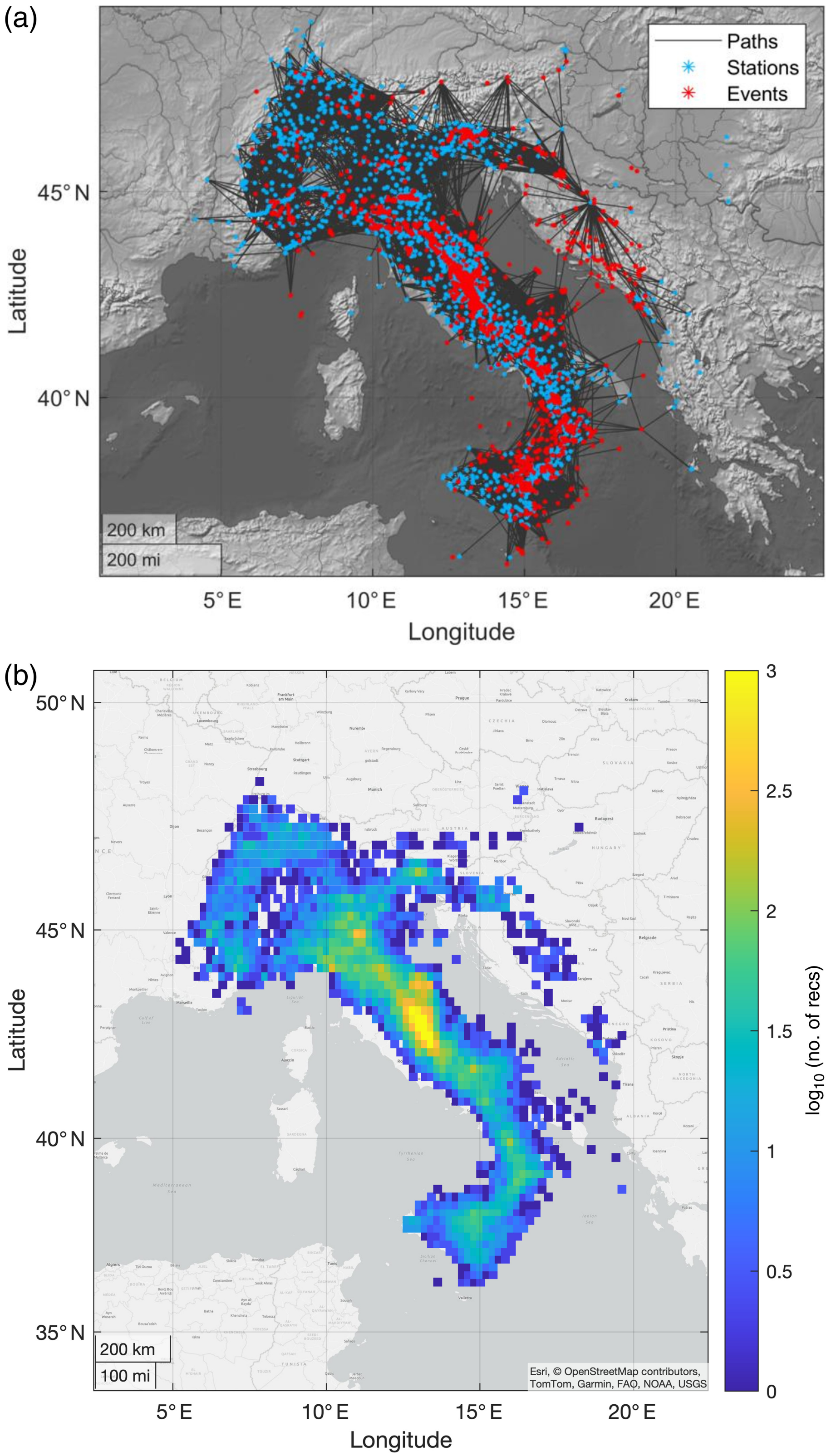


Figure 9. Path coverage of ITACAext_sac. (a) Source to site paths; (b) source to site middle point paths density measured considering a 0.2° × 0.2° grid. The color version of this figure is available only in the electronic edition.

CMB station, we cannot attribute the peculiar shape of the site amplification to local site effects, housing effects, or an instrumental problem. The site amplifications in Figure 10 show a decay at high frequencies with variable slope for the different stations, consistent with the empirical model describing the near-surface attenuation by the exponential function $A(f) = \exp(-\pi k_0 f)$ (Anderson and Hough, 1984). In summary, the GIT results confirm that the quality of the Fourier spectra included in ITACAex_sac is suitable for the study of source scaling, propagation effects, and site amplifications in Italy through spectral decomposition approaches. For the sake of completeness, the supplemental appendices include the other two GIT terms obtained from the spectral decomposition. The spectral attenuation with distance and the source spectra shown in Figure A5 confirm that the GIT results are robust and consistent with previous studies (e.g., Bindi and Kotha, 2020).

Flatfile Dissemination

The ITACAext flatfiles (v.1.0 and 2.0) are accessible through the ITACA database website (see Data and Resources). Ground-motion amplitudes and event, station, and waveform metadata are distributed through the standard format of the ESM flatfile, which is described in detail in Lanzano, Sgobba, *et al.* (2019). Briefly, the fields of the flatfile can be grouped into six main blocks: (1) event-related metadata; (2) source-related metadata; (3) station-related metadata;

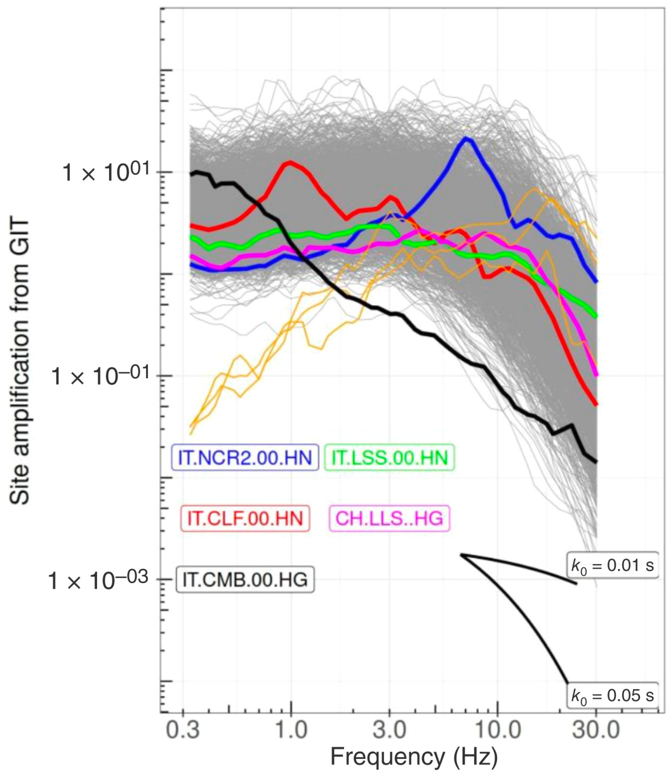


Figure 10. Site amplifications (gray lines) for stations in ITACAext_sac obtained by generalized inversion technique (GIT) spectral decomposition. The results for a sample of selected stations (names given in the panel) are shown in color; the orange lines indicate the results for three stations installed in the Balkans. The two black lines in the lower right corner show the exponential trend of the function $\exp(-\pi k_0 f)$ with $k_0 = 0.01$ and 0.05 s, which is provided as a reference for the high-frequency decay of the site amplifications. The color version of this figure is available only in the electronic edition.

(4) metrics of source-to-site distances; (5) waveform-related metadata; and (6) intensity measures (IMs). In particular, the ITACAext 2.0 release consists of four tables in csv format (semicolon separated):

- ITACAext_flatfile_2_SA.csv: a parametric table containing the SA ordinates (5% damping) and the metadata necessary to identify each waveform of ITACAext 2.0 dataset (event_id; network_code; station_code; location_code; instrument_code);
- ITACAext_flatfile_2_FAS.csv: a parametric table containing the FAS ordinates (smoothing Konno and Omachi, 1998, $b = 40$) and the metadata necessary to identify each waveform of ITACAext 2.0 dataset (event_id; network_code; station_code; location_code; instrument_code;);
- ITACAext_flatfile_2_FAS_S.csv: a parametric table containing the FAS ordinates computed on the *S* phase (smoothing Konno and Omachi, 1998, $b = 40$) and the metadata necessary to identify each waveform of ITACAext 2.0 dataset (event_id; network_code; station_code; location_code; instrument_code;); and
- ITACAext_2_flatfile_METADATA.csv: a parametric table containing all the available metadata for each waveform in the ITACAext 2.0 dataset.

The tables are included in a zip file with a user manual (User_manual_ITACAext_2.pdf, also available in ESUPP2) and several dictionaries related to some fields of the metadata table. The products are citable using a digital object identifier (DOI), with the full citation given in the reference list as Lanzano *et al.* (2024).

The ITACAext strong motion flatfiles are licensed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License. This means that the user is free to share (copy and redistribute the material in any medium or format) and adapt (remix, transform, and build upon the material) for any purpose.

Discussions and Conclusions

The construction of qualified datasets of strong and weak motion seismic records is a key step in calibrating and testing ground-motion prediction models for seismological and engineering purposes. In the past 20 yr, many efforts have been made to build good-quality datasets by collecting strong earthquake observations at global or continental scales. More recently, the extraordinary growth of seismic networks in some countries and the occurrence of significant seismic sequences (in terms of data collected) have pushed the construction of regional datasets including lower-magnitude data as well.

Following the example of datasets produced for other countries, we decided to replicate the effort by building a high-quality dataset for Italy. It was decided to exploit the great legacy available in the ITACA database not only for the large number of manually processed waveforms of weak-to-strong earthquakes since the 1970s but also for the wealth of quality metadata associated with both events and recording stations. The new flatfile is called ITACAext, intended as an extended version of the ITACA database because we decided to also include the velocimetric recordings and data from seismic networks of neighboring countries, with the aim of improving the description of wave propagation characteristics at regional/local scale. ITACAext is composed of data and metadata of 52,706 recordings relative to 2,338 events (until December 2022) and 2,294 stations. The proposed dataset includes records of earthquakes from different tectonic contexts such as active shallow crustal areas (in most cases), volcanic districts (Mount Etna, Campi Flegrei, Ischia, Vesuvio, Colli Albani, and Aeolian Islands), subduction zones (in the lower Tyrrhenian Sea and Calabria), and deep seismicity (in the northern Apennines). Furthermore, the inclusion of small-magnitude data and velocimeteric recordings makes the dataset particularly suitable for studying differences in ground motion in different parts of Italy, in line with the study by Brunelli *et al.* (2023).

For each recordings, we computed the main relevant engineering parameters, such as peak parameters (PGA, PGV, and PGD), energy content parameters (AI, Housner intensity, and CAV), duration (significant duration D5–95), the ordinates of acceleration response spectra (5% damping) in the interval 0.01–10 s, and the ordinates of the FAS from 0.1 to 30 Hz.

In view of the hazard applications, a further analysis is performed to construct an optimized best-quality subset to study the seismicity characteristics of active crustal events, the most representative type of seismicity across Italy. The selection procedure involves two stages (Table 1).

In the first stage, basic filters are applied to the dataset based on magnitude, distance, depth, and geographic coordinates (to exclude surface events in volcanic districts).

In the second stage, the quality of the dataset is analyzed using a two-step residual analysis. In the first step, by means of residuals decomposition of the FAS ordinates, the site responses of colocated sensors are compared, with the goal of identifying anomalous behaviors of individual sensors and making a choice between colocated sensor records when both are available for the same event. The comparison shows a systematic difference between accelerometric and velocimetric sensors only at high frequencies ($f > 15$ Hz) and, to a lesser extent, at low frequencies ($f < 0.2$ Hz). The observed differences, especially at high frequencies, could be related to the different installation conditions of the accelerometer and velocimeter instruments (Hollender *et al.*, 2020). This inconsistency makes it difficult to analyze the amplitude of the Fourier spectra at high frequencies ($f > 10$ Hz) and the calculation of parameters for this frequency range, such as the κ_0 parameter, that is, the near-site attenuation at high frequencies (Anderson and Hough, 1984; Ktenidou *et al.*, 2014). The second residual step is a standard consistency check obtained by calibrating a partially nonergodic GMM over the data and identifying outliers in the distributions of events, stations, and aleatory residuals for some ground-motion parameters of engineering interest.

The analysis resulted in the selection of a good-quality subset for the study of crustal active ground motion in Italy, consisting of 42,877 records, for 2,052 events, and 2,072 stations, corresponding to ~81% of the original dataset. The selected data are easily identifiable in the table by means of a flag field in the ITACAext 2.0 flatfile; the discarded data, on the other hand, were not removed from the table but were listed in the appendices for later analysis and verification.

As a final application, we conducted a GIT analysis following the procedure of Bindi and Kotha (2020) focusing on site functions: the seismic response of some well-studied sites is compatible with previous studies and seismological analyses using different datasets and methods. However, some anomalous sites are still present in the dataset (such as the case of the CMB station), which deserve further investigation.

Finally, it should be noted that the approach taken to identify the high-quality subset is conservative, discarding uncertain data and ensuring that only significant outliers that cannot be attributed to peculiar ground-motion characteristics (near-source effects, directivity, basin effects, etc.) are removed. The applied procedure also allows for (partial) automation, especially in the identification of outliers from the residual analysis and would allow the flatfile to be updated regularly.

Data and Resources

The.zip files of ITACAext flatfiles (v.1.0 and 2.0) are accessible at the permanent URL https://itaca.mi.ingv.it/ItacaNet_40/#/products/itacaext_flatfile. The accelerometric data used in ITACAext are all accessible through the ITalian ACcelerometric Archive (ITACA, https://itaca.mi.ingv.it/) and Engineering Strong Motion Database (ESM, https://esm-db.eu/) websites. Velocimeter records were extracted from the continuous streaming of stations available in European Integrated Data Archive (EIDA) at https://www.orfeus-eu.org/data/eida/. More information regarding contributing networks can be found in ITACA website at https://itaca.mi.ingv.it/ItacaNet_40/#/about/who/contributing_networks or on the International Federation of Digital Seismograph Networks (FDSN) website (https://www.fdsn.org/). All websites were last accessed in May 2025. The supplemental material includes some additional figures described in the text (ESUPP1); the ITACAext user manual (ESUPP2); and a list of the outliers, identified after the quality check (ESUPP3).

Declaration of Competing Interests

The authors acknowledge that there are no conflicts of interest recorded.

Acknowledgments

The ITACAext flatfile has been developed within the INGV-DPC (Istituto Nazionale di Geofisica e Vulcanologia-Civil Protection Department) Agreement (2022–2024 and 2024–2026)—Annex A, in the framework of the activities of ITalian ACcelerometric Archive (ITACA) database in the Task 7.2 (Seismological instrumental databases). The product development is also supported by the INGV CPS (Seismic Hazard Center) and the research project INGV-MUR Working Earth: geosciences and understanding of the earth dynamics and natural hazards, in the framework of SECURE (Regional-scale earthquake ground-motion predictions through physics-based and empirical approaches: A case study central Italy) project. Some of the analyses presented in this work are also the result of discussions in the Strong Motion Committee (SMC-SM) of Observatories and Research Facilities for European Seismology (ORFEUS) regarding the comparison of records from colocated sensors. The authors wish also to thank the reviewers of the article (Lisa Schleicher and an anonymous reviewer) they provided helpful comments.

References

Akkar, S., Z. Çağnan, E. Yenier, Ö. Erdoğan, M. A. Sandıkkaya, and P. Gülkan (2010). The recently compiled Turkish strong motion database: Preliminary investigation for seismological parameters, *J. Seismol.* **14,** 457–479.

Akkar, S., M. A. Sandıkkaya, M. Şenyurt, A. A. Sisi, B. Ö. Ay, P. Traversa, J. Douglas, F. Cotton, L. Luzi, B. Hernandez, and S. Godey (2014). Reference database for seismic ground-motion in Europe (RESORCE), *Bull. Earthq. Eng.* **12,** no. 1, 311–339.

Al Atik, L., N. Abrahamson, J. J. Bommer, F. Scherbaum, F. Cotton, and N. Kuehn (2010). The variability of ground-motion prediction models and its components, *Seismol. Res. Lett.* **81,** no. 5, 794–801, doi: 10.1785/gssrl.81.5.794.

Alparone, S., M. Cattaneo, D. Contrafatto, E. D'Alema, D. Franceschi, G. Larocca, A. G. Mandiello, S. Marzorati, M. Moretti, D. Piccinini, *et al.* (2020). *Seismic Data acquired by the SISMIKO Emergency Group - Salemi-Italy 2020 - T15* [Data set], Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.13127/sd/ljfpeosek0.

Alparone, S., O. M. G. Cocina, D. Contrafatto, G. Larocca, M. Manni, L. Pruiti, A. A. Messina, A. Ursino, and E. D'Alema (2021). *Seismic Data acquired by the INGV temporary network for seismic monitoring of the island of Vulcano to increase volcanic activity (Sicily, Italy) - 2021* [Data set], Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.13127/sd/cfz7ftsw-y.

AlpArray Seismic Network (2015). AlpArray Seismic Network (AASN) temporary component, AlpArray Working Group, doi: 10.12686/alparray/z3_2015.

Ambraseys, N., P. Smit, J. Douglas, B. Margaris, R. Sigbjornsson, S. Olafsson, P. Suhadolc, and G. Costa (2004). Internet-site for European strong motion data, *Boll. Geofis. Teor. Appl.* **45,** no. 3, 113–129.

Ambraseys, N. N., J. Douglas, S. K. Sarma, and P. M. Smit (2005). Equations for the estimation of strong ground motions from shallow crustal earthquakes using data from Europe and the Middle East: Horizontal peak ground acceleration and spectral acceleration, *Bull. Earthq. Eng.* **3,** 1–53.

Ancheta, T. D., R. B. Darragh, J. P. Stewart, E. Seyhan, W. J. Silva, B. S. J. Chiou, K. E. Wooddell, R. W. Graves, A. R. Kottke, D. M. Boore, and T. Kishida (2014). NGA-West2 database, *Earthq. Spectra* **30,** no. 3, 989–1005.

Anderson, J. G., and S. E. Hough (1984). A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies, *Bull. Seismol. Soc. Am.* **74,** no. 5, 1969–1993.

Art Lerner-Lam (2003). *Calabria-Appennine-Tyrrhenian/Subduction-Collision-Accretion Network* [Data set], International Federation of Digital Seismograph Networks, doi: 10.7914/SN/YD_2003.

Bahrampouri, M., A. Rodriguez-Marek, S. Shahi, and H. Dawood (2021). An updated database for ground motion parameters for KiK-net records, *Earthq. Spectra* **37,** no. 1, 505–522.

Bastías, N., and G. A. Montalva (2016). Chile strong ground motion flatfile, *Earthq. Spectra* **32,** no. 4, 2549–2566.

Bates, D., M. Mächler, B. Bolker, and S. Walker (2015). Fitting linear mixed-effects models using lme4, *J. Stat. Software* **67,** no. 1, 1–48.

Bindi, D., and S. R. Kotha (2020). Spectral decomposition of the Engineering strong motion (ESM) flat file: Regional attenuation, source scaling and Arias stress drop, *Bull. Earthq. Eng.* doi: 10.1007/s10518-020-00796-1.

Bindi, D., S. R. Kotha, G. Weatherill, G. Lanzano, L. Luzi, and F. Cotton (2019). The pan-European engineering strong motion (ESM) flatfile: Consistency check via residual analysis, *Bull. Earthq. Eng.* **17,** 583–602.

Bindi, D., R. Zaccarelli, H. N. T. Razafindrakoto, M.-H. Yen, and F. Cotton (2023). Empirical shaking scenarios for Europe: A feasibility study, *Geophys. J. Int.* **232,** no. 2, 990–1005, doi: 10.1093/gji/ggac382.

Bommer, J. J., P. J. Stafford, and J. E. Alarcón (2009). Empirical equations for the prediction of the significant, bracketed, and uniform duration of earthquake ground motion, *Bull. Seismol. Soc. Am.* **99,** no. 6, 3217–3233.

Boore, D. M. (2010). Orientation-independent, nongeometric-mean measures of seismic intensity from two horizontal components of motion, *Bull. Seismol. Soc. Am.* **100,** no. 4, 1830–1835.

Brunelli, G., G. Lanzano, M. C. D'Amico, C. Felicetta, L. Luzi, C. Mascandola, F. Pacor, E. Russo, and S. Sgobba (2022). *ITACAext flatfile* [Data set], Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.13127/itaca32/itacaext_flatfile.1.0.

Brunelli, G., G. Lanzano, L. Luzi, and S. Sgobba (2023). Data-driven zonations for modelling the regional source and propagation effects into ground motion models in Italy, *Soil Dynam. Earthq. Eng.* **166,** 107775, doi: 10.1016/j.soildyn.2023.107775.

Cara, F., G. Cultrera, G. Riccio, S. Amoroso, P. Bordoni, A. Bucci, E. D'Alema, M. D'Amico, L. Cantore, S. Carannante, *et al.* (2019). Temporary dense seismic network during the 2016 central Italy seismic emergency for microzonation studies, *Scientific Data* **6,** no. 1, 182.

Cauzzi, C., D. Bindi, D. Cambaz, F. Carrilho, S. Custódio, J. Clinton, C. Felicetta, C. Evangelidis, F. Gallovič, D. Kalafat, *et al.* (2022). Systematic comparisons of broad-band velocity and acceleration earthquake records as a quality assessment tool for European open strong-motion data, *Seismol. Res. Lett.* **93,** no. 2B, 1115–1372.

CERN (2016). CERN Seismic Network, *ETH Zurich*, doi: 10.12686/sed/networks/c4.

Chiarabba, C., A. Costanzo, P. De Gori, P. Del Gaudio, S. Falcone, R. Fonzetti, A. Gervasi, A. Govoni, C. La Piana, F. P. Lucente, *et al.* (2022). *Pollino Broad-Band Near Fault Observatory (2022-2033), Southern Italy* [Data set], Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.13127/sd/ns6hb3q8ve.

Chiou, B., R. Darragh, N. Gregor, and W. Silva (2008). NGA project strong-motion database, *Earthq. Spectra* **24,** no. 1, 23–44.

Cornou, C., P.-Y. Bard, and , and RESIF (2014). *Seismic network YP: French part of the Argostoli (Greece) aftershock experiment (RESIF-SISMOB)* [Data set], RESIF - Réseau Sismologique et géodésique Français, doi: 10.15778/RESIF.YP2014.

Costanzo, A., S. Falcone, C. La Piana, A. Caserta, F. Doumaz, M. F. Buongiorno, A. D'Alessandro, G. Vitale, S. Scudero, and S. Speciale (2022). *Urban Seismic Network (RUSN-CAL-INGV) - Real time Urban Seismic Network managed by CALabrian headquarters of INGV for monitoring urban areas, critical infrastructures and cultural heritage sites* [Data set], Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.13127/sd/o5nrhmp9n7.

D'Alema, E., S. Alparone, P. Augliera, D. Biagini, C. Calamita, A. Castagnozzi, A. Cavaliere, A. Costanzo, E. Della Bina, S. Farroni, *et al.* (2022). *Seismic Data acquired by the SISMIKO Emergency Group - Northern Marche Coast - Italy 2022 - T17* [Data set], Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.13127/sd/tblkba-3u6.

Dawood, H. M., A. Rodriguez-Marek, J. Bayless, C. Goulet, and E. Thompson (2016). A flatfile for the KiK-net database processed using an automated protocol, *Earth. Spectra* **32,** no. 2, 1281–1302.

Department of Earth and Environmental Sciences, Geophysical Observatory, University of Munchen (LMU) (2001). *BayernNetz* [Data set], International Federation of Digital Seismograph Networks, doi: 10.7914/SN/BW.

Di Giulio, G., A. Rovelli, F. Cara, R. M. Azzara, F. Marra, R. Basili, and A. Caserta (2003). Long-duration asynchronous ground motions in the Colfiorito plain, central Italy, observed on a two-dimensional dense array, *J. Geophys. Res.* **108,** no. B10, 2486, doi: 10.1029/2002jb002367.

Dreger, D. S., W. H. K. Lee, H. Kanamori, P. C. Jennings, and C. Kisslinger (2003). Time-domain moment tensor INVerse codel (TDMT-INVC) release 1.1, *Int. Handb. Earthq. Eng. Seismol* **81,** 1627.

Douglas, J., and D. M. Boore (2011). High-frequency filtering of strong-motion records, *Bull. Earthq. Eng.* **9,** 395–409.

EMERSITO Working Group (2018). Rete sismica del gruppo EMERSITO, sequenza sismica del 2016 in Italia Centrale, Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.13127/SD/7TXeGdo5X8 (in Italian).

EMERSITO Working Group (2024). *Rete sismica temporanea 6N installata ad Ancona dal gruppo EMERSITO durante la sequenza sismica del 2022 - Costa Marchigiana-Pesarese* [Data set], Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.13127/sd/qctgd6c-3a (in Italian).

Epos-France Seismology (RLBP) (1995). *Epos-France Broad-band network (RLBP)* [Data set], Epos-France Seismological Data Centre, doi: 10.15778/RESIF.FR.

European Committee for Standardization (CEN) (2003). *Eurocode 8. Design of Structures for Earthquake Resistance—Part 1: General Rules Seismic Actions and Rules for Buildings, EN 1998-1*, European Committee for Standardization, Brussels, Belgium.

Felicetta, C., E. Russo, M. D'Amico, S. Sgobba, G. Lanzano, C. Mascandola, F. Pacor, and L. Luzi (2023). *Italian Accelerometric Archive v4.0 - Istituto Nazionale di Geofisica e Vulcanologia, Dipartimento della Protezione Civile Nazionale*, doi: 10.13127/itaca.4.0 (in Italian).

Gallipoli, M. (2003). *Temporary Accelerometric Network in Basilicata region for Site Effect studies* [Data set], International Federation of Digital Seismograph Networks, doi: 10.7914/nhc9-gg46.

Galluzzo, D., L. Nardone, G. Gaudiosi, C. Martino, G. Scarpato, A. Carandente, C. Buonocunto, A. Govoni, M. Moretti, E. D'Alema, and , and SISMIKO Working Group (2017). *Seismic Data acquired by the SISMIKO Emergency Group - Ischia-Italy 2017 - T13* [Data set], Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.13127/sd/u0j8hsaevf.

Gasperini, P., B. Lolli, and G. Vannucci (2013). Empirical calibration of local magnitude data sets versus moment magnitude in Italy, *Bull. Seismol. Soc. Am.* **103,** no. 4, 2227–2246.

Gasperini, P., B. Lolli, G. Vannucci, and E. Boschi (2012). Calibration of moment magnitude estimates for the European–Mediterranean and Italian regions, *Geophys. J. Int.* **190,** no. 3, 1733–1745.

GEOFON Data Centre (1993). *GEOFON Seismic Network* [Data set], GFZ Data Services, doi: 10.14470/TR560404.

Geological Survey-Provincia Autonoma di Trento (1981). *Trentino Seismic Network* [Data set], International Federation of Digital Seismograph Networks, doi: 10.7914/SN/ST.

Grünthal, G., and R. Wahlström (2012). The European-Mediterranean earthquake catalogue (EMEC) for the last millennium, *J. Seismol.* **16,** no. 3, 535–570.

Helmstetter, A., P. Guéguen, and , and RESIF (2020). *Seismic network 8C: Monitoring swarms in the Mont-Blanc and Vallorcine area (RESIF - SISMOB)* [Data set], RESIF - Réseau Sismologique et géodésique Français, doi: 10.15778/RESIF.8C2019.

Hollender, F., Z. Roumelioti, E. Maufroy, P. Traversa, and A. Mariscal (2020). Can we trust high-frequency content in strong-motion database signals? Impact of housing, coupling, and installation depth of seismic sensors, *Seismol. Res. Lett.* **91,** no. 4, 2192–2205.

Hutchinson, J. A., C. Zhu, B. A. Bradley, R. L. Lee, L. M. Wotherspoon, M. Dupuis, C. Schill, J. Motha, E. Manea, and A. E. Kaiser (2024). The 2023 New Zealand ground-motion database, *Bull. Seismol. Soc. Am.* **114,** no. 1, 291–310.

Institut de physique du globe de Paris (IPGP) and École et Observatoire des Sciences de la Terre de Strasbourg (EOST) (1982). *GEOSCOPE, French Global Network of broad band seismic stations*, Institut de physique du globe de Paris (IPGP), Université de Paris, doi: 10.18715/GEOSCOPE.G.

Institute of GeoSciences, Polytechnic University of Tirana (IGEO PUT) (2002). *Albanian Seismological Network* [Data set], International Federation of Digital Seismograph Networks, doi: 10.7914/SN/AC.

Istituto Nazionale di Geofisica e Vulcanologia (INGV) (2005). *Rete Sismica Nazionale (RSN) [Data set]. Istituto Nazionale di Geofisica e Vulcanologia (INGV)*, doi: 10.13127/sd/x0fxnh7qfy.

INGV (2009). Emersito Seismic Network for Site Effect Studies in L'Aquila town (Central Italy), International Federation of Digital Seismograph Networks, doi: 10.7914/SN/4A_2009 (in Italian).

INGV, Istituto di Geologia Ambientale e Geoingegneria (CNR-IGAG), Istituto per la Dinamica dei Processi Ambientali (CNR-IDPA), Istituto di Metodologie per l'Analisi Ambientale (CNR-IMAA), & Agenzia Nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile (ENEA) (2018). *Rete del Centro di Microzonazione Sismica (CentroMZ), sequenza sismica del 2016 in Italia Centrale* [Data set], Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.13127/sd/ku7xm12yy9 (in Italian).

International Seismological Centre (ISC) (2020). International Seismological Centre: On-line Bulletin, doi: 10.31905/D808B830.

ISIDe Working Group (2007). *Italian Seismological Instrumental and Parametric Database (ISIDe)*, Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.13127/ISIDE.

Istituto Nazionale di Oceanografia e di Geofisica Sperimentale - OGS (2016). *North-East Italy Seismic Network* [Data set], FDSN, doi: 10.7914/SN/OX.

Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (OGS) and University of Trieste (2002). North-East Italy Broadband Network [Data set], International Federation of Digital Seismograph Networks, doi: 10.7914/SN/NI.

Konno, K., and T. Ohmachi (1998). Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor, *Bull. Seismol. Soc. Am.* **88,** no. 1, 228–241.

Kotha, S. R., G. Weatherill, D. Bindi, and F. Cotton (2020). A regionally-adaptable ground-motion model for shallow crustal earthquakes in Europe, *Bull. Earthq. Eng.* **18**, no. 9, 4091–4125.

Ktenidou, O. J., F. Cotton, N. A. Abrahamson, and J. G. Anderson (2014). Taxonomy of κ: A review of definitions and estimation approaches targeted to applications, *Seismol. Res. Lett.* **85,** no. 1, 135–146.

Lammers, S., G. Weatherill, G. Grünthal, and F. Cotton (2023). EMEC-2021–The European-Mediterranean Earthquake Catalogue–Version 2021, GFZ German Research Center for Geosciences, doi: 10.5880/GFZ.EMEC.2021.001.

Lanzano, G., C. Felicetta, F. Pacor, D. Spallarossa, and P. Traversa (2022). Generic-to-reference rock scaling factors for seismic ground motion in Italy, *Bull. Seismol. Soc. Am.* **112,** no. 3, 1583–1606.

Lanzano, G., L. Luzi, F. Pacor, C. Felicetta, R. Puglia, S. Sgobba, and M. D'Amico (2019). A revised ground-motion prediction model for shallow crustal earthquakes in Italy, *Bull. Seismol. Soc. Am.* **109,** no. 2, 525–540.

Lanzano, G., S. Sgobba, L. Luzi, R. Puglia, F. Pacor, C. Felicetta, M. D'Amico, F. Cotton, and D. Bindi (2019). The pan-European Engineering Strong Motion (ESM) flatfile: Compilation criteria and data statistics, *Bull. Earthq. Eng.* **17,** 561–582.

Lanzano, G., L. Vitrano, C. Felicetta, E. Russo, M. C. D' Amico, C. Mascandola, S. Sgobba, G. Brunelli, F. Ramadan, F. Pacor, and L. Luzi (2024). *ITACAext flatfiles: parametric tables of metadata and strong motion intensity measures (Version 2.0)* [Data set], Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.13127/itaca.4.0/itacas2s_flatfile.2.0.

Luzi, L., G. Lanzano, C. Felicetta, M. C. D'Amico, E. Russo, S. Sgobba, F. Pacor, and ORFEUS Working Group 5 (2020). Engineering Strong Motion Database (ESM) (Version 2.0), Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.13127/ESM.2.

Margaris, B., E. M. Scordilis, J. P. Stewart, D. M. Boore, N. Theodoulidis, N. I. Kalogeras, N. S. Melis, A. A. Skarlatoudis, N. Klimis, and E. Seyhan (2021). Hellenic Strong-Motion Database with Uniformly Assigned Source and Site Metadata for the Period 1972–2015, *Seismol. Res. Lett.* **92,** no. 3, 2065–2080.

Marzorati, S., C. Ladina, D. Pantaleo, C. Calamita, S. Sofia, F. Boccanera, and G. Speranza (2021). *Seismic Data acquired by Snowbilla network - Sibillini Mountains - Italy* [Data set], Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.13127/sd/fcjdr8key3.

Mascandola, C., G. Lanzano, and F. Pacor (2022). Consistency Check of ITACAext, the Flatfile of the Italian Accelerometric Archive, *Geosciences* **12,** no. 9, 334.

Mascandola, C., L. Luzi, C. Felicetta, and F. Pacor (2021). A GIS procedure for the topographic classification of Italy, according to the seismic code provisions, *Soil Dynam. Earthq. Eng* **148,** 106848.

Massa, M., A. L. Rizzo, A. Lorenzetti, S. Lovati, E. D'Alema, R. Puglia, S. Carannante, and L. Luzi (2021). *Rete di monitoraggio multiparametrico del Garda (Nord Italia) - PDnet (Version 1.0)* [Data set], Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.13127/sd/yhcfomcbo_.

MedNet Project Partner Institutions (1990). Mediterranean Very Broadband Seismographic *Network*(MedNet), Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.13127/SD/fBBBtDtd6q.

Morasca, P., D. Bindi, K. Mayeda, J. Roman-Nieves, J. Barno, W. R. Walter, and D. Spallarossa (2022). Source scaling comparison and validation in central Italy: Data intensive direct s waves versus the sparse data coda envelope methodology, *Geophys. J. Int.* **231,** 1573–1590.

Morasca, P., M. D'Amico, S. Sgobba, G. Lanzano, L. Colavitti, F. Pacor, and D. Spallarossa (2023). Empirical correlations between an FAS non-ergodic ground motion model and a GIT derived model for Central Italy, *Geophys. J. Int.* **233,** no. 1, 51–68.

Moretti, M., P. De Gori, A. Govoni, L. Margheriti, D. Piccinini, S. Pintore, and L. Valoroso (2018). Seismic Data acquired by the SISMIKO Emergency Group - Molise-Italy 2018 - T14 [Data set], Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.13127/sd/fir72chywu.

Moretti, M., L. Margheriti, S. Alparone, A. Costanzo, M. La Rocca, S. Murphy, M.-A. Gutscher, and Focus Working Group. (2021). Seismic Data acquired by FocusX temporary land-network (FXland), Southern Italy [Data set], Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.13127/sd/o5qwm6wjcd.

National Institute for Earth Physics (NIEP Romania) (1994). Romanian Seismic Network [Data set], International Federation of Digital Seismograph Networks, doi: 10.7914/SN/RO.

National Observatory of Athens (NOA), Institute of Geodynamics, Athens (1975). National Observatory of Athens Seismic Network [Data set], International Federation of Digital Seismograph Networks, doi: 10.7914/SN/HL.

Pacor, F., R. Paolucci, L. Luzi, F. Sabetta, A. Spinelli, A. Gorini, M. Nicoletti, S. Marcucci, L. Flippi, and M. Dolce (2011). Overview of the Italian strong motion database ITACA 1.0, *Bull. Earthq. Eng.* **9,** 1723–1739.

Paolucci, R., F. Pacor, R. Puglia, G. Ameri, C. Cauzzi, and M. Massa (2011). Record processing in ITACA, the new Italian strong-motion database, earthquake data, in *Engineering Seismology*, S. Akkar, P. Gülkan, and T. Van Eck (Editors), Springer, Dordrecht, The Netherlands, 99–113.

Piccinini, D., D. Biagini, M. D'Ambrosio, E. D'Alema, L. Zuccarello, G. Saccorotti, M. Pastori, I. Carluccio, M. Fares, E. Della Bina, M. Quintiliani, *et al.* (2022). *Seismic Data acquired by the SISMIKO Emergency Group - Chianti Fiorentino - Italy 2022 - T16* [Data set], Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.13127/sd/jhggg3bnlq.

Pilz, M., and S. Parolai (2009). *Norcia basin (Italy) temporary seismic network* [Data set], GFZ Data Services, doi: 10.14470/8U7554472182.

Pondrelli, S., A. Morelli, G. Ekström, S. Mazza, E. Boschi, and A. M. Dziewonski (2002). European–Mediterranean regional centroid-moment tensors: 1997–2000, *Phys. Earth Planet In.* **130,** nos. 1/2, 71–101.

Pondrelli, S., F. Visini, A. Rovida, V. D'Amico, B. Pace, and C. Meletti (2020). Style of faulting of expected earthquakes in Italy as an input for seismic hazard modeling, *Nat. Hazards Earth Syst. Sci.* **20,** no. 12, 3577–3592.

Presidency of Council of Ministers - Civil Protection Department (DPC) (1972). *Italian strong motion network* [Data set], International Federation of Digital Seismograph Networks, doi: 10.7914/SN/IT.

Presidenza del Consiglio dei Ministri - Dipartimento della protezione civile (RACSI) (2017). *Reti accelerometriche centri storici italiani* [Data set], International Federation of Digital Seismograph Networks, doi: 10.7914/SN/RR.

Puglia, R., E. Russo, L. Luzi, M. D'Amico, C. Felicetta, F. Pacor, and G. Lanzano (2018). Strong-motion processing service: A tool to access and analyse earthquakes strong-motion waveforms, *Bull. Earthq. Eng.* **16,** 2641–2651.

RESIF (1995). *RESIF-RAP French Accelerometric Network* [Data set], RESIF - Réseau Sismologique et géodésique Français, doi: 10.15778/RESIF.RA.

Rovelli, A., A. Caserta, F. Marra, and V. Ruggiero (2002). Can seismic waves be trapped inside an inactive fault zone? The case study of Nocera Umbra, central Italy, *Bull. Seismol. Soc. Am.* **92,** 2217–2232.

Sector for Seismology, Institute of Hydrometeorology and Seismology of Montenegro (ME) (1982). *Montenegrin Seismic Network* [Data set], International Federation of Digital Seismograph Networks, doi: 10.7914/SN/ME.

Slovenian Environment Agency (1990). *Seismic Network of the Republic of Slovenia* [Data set], International Federation of Digital Seismograph Networks, doi: 10.7914/SN/SL.

Spallarossa, D., M. Cattaneo, D. Scafidi, M. Michele, L. Chiaraluce, M. Segou, and I. G. Main (2021). An automatically generated high-resolution earthquake catalogue for the 2016–2017 central Italy seismic sequence, including P and S phase arrival times, *Geophys. J. Int.* **225,** no. 1, 555–571.

Strollo, A., D. Cambaz, J. Clinton, P. Danecek, C. P. Evangelidis, A. Marmureanu, and N. Triantafyllis (2021). EIDA: The European integrated data archive and service infrastructure within ORFEUS, *Bull. Seismol. Soc. Am.* **92,** no. 3, 1788–1795.

Swiss Seismological Service (SED) At ETH Zurich (1983). National Seismic Networks of Switzerland, ETH Zürich, doi: 10.12686/sed/networks/ch.

Traversa, P., E. Maufroy, F. Hollender, V. Perron, V. Bremaud, H. Shible, S. Drouet, M. Langlais, D. Wolyniec, C. Péquegnat, and I. Douste-Bacque (2020). RESIF RAP and RLBP dataset of earthquake ground motion in mainland France, *Seismol. Res. Lett.* **91,** no. 4, 2409–2424.

Trifunac, M. D., and A. G. Brady (1975). A study on the duration of strong earthquake ground motion, *Bull. Seismol. Soc. Am.* **65,** no. 3, 581–626.

Universita Della Calabria, Italy (Unical) (1981). *Rete Sismica Unical* [Data set], International Federation of Digital Seismograph Networks, doi: 10.7914/SN/IY.

University of Genoa (UniGE) (1967). *Regional Seismic Network of North Western Italy* [Data set], International Federation of Digital Seismograph Networks, doi: 10.7914/SN/GU.

University of Trieste (1993). *Friuli Venezia Giulia Accelerometric Network* [Data set],International Federation of Digital Seismograph Networks, doi: 10.7914/SN/RF.

Vadim, L. (2003). *Retreating-Trench, Extension, and Accretion LTectonics: A Multidisciplinary Study of the Northern Apennines* [Data set], International Federation of Digital Seismograph Networks, doi: 10.7914/SN/YI_2003.

Visini, F., C. Meletti, A. Rovida, V. D'Amico, B. Pace, and S. Pondrelli (2022). An updated area-source seismogenic model (MA4) for seismic hazard of Italy, *Nat. Hazards Earth Syst. Sci.* **22,** no. 8, 2807–2827.

Vitrano, L., L. Luzi, J. Clinton, M. C. D'Amico, C. Mascandola, C. Felicetta, E. Russo, G. Lanzano, S. Sgobba, and C. Cauzzi (2024). ESM processing tool users' manual, *Rapporti Tecnici INGV n°477*.

Wald, D. J., and T. I. Allen (2007). Topographic slope as a proxy for seismic site conditions and amplification, *Bull. Seismol. Soc. Am.* **97,** no. 5, 1379–1395.

ZAMG - Zentralanstalt für Meterologie und Geodynamik (1987). *Austrian Seismic Network* [Data set], International Federation of Digital Seismograph Networks, doi: 10.7914/SN/OE.

Zare, M., and P. Y. Bard (2002). Strong motion dataset of Turkey: Data processing and site classification, *Soil Dynam. Earthq. Eng.* **22,** no. 8, 703–718.

Zhao, L., A. Paul, S. Solarino, and , and RESIF (2016). *Seismic network YP: CIFALPS temporary experiment (China-Italy-France Alps seismic transect)* [Data set], RESIF - Réseau Sismologique et géodésique Français, doi: 10.15778/RESIF.YP2012.

Appendix A

List of the contributing networks

This appendix reports a table (Table A1) with the list of the contributing networks to ITACAext 2.0, including some brief info.

TABLE A1
List of the Contributing Networks to ITACAext 2.0

Network Code	No. of Data	%	Name/Description	Abbreviated Reference*	DOI
1J	20	<0.1	FocusX temporary land network (FXland)	Moretti *et al.* (2021)	10.13127/sd/o5qwm6wjcd
3A	1537	2.9	Centro di microzonazione sismica Network, 2016 central Italy seismic sequence (CentroMZ)	INGV *et al.* (2018)	10.13127/sd/ku7xm12yy9
3D	10	<0.1	Istituto Nazionale di Geofisica e Vulcanologia (INGV) temporary network for seismic monitoring of the island of Vulcano to increase volcanic activity (Sicily, Italy) (VITN)	Alparone *et al.* (2021)	10.13127/sd/cfz7ftsw-y
3H	1075	2	Norcia basin (Italy) temporary network	Pilz and Parolai (2009)	10.14470/8U7554472182
4A	479	0.9	Emersito Seismic Network for site effect studies in L'Aquila town (central Italy) (ESSEA)	INGV (2009)	10.7914/SN/4A_2009

*The full citation is reported in the reference list.
†The network code is now changed to 1V (Gallipoli, 2003). Full citation in the reference list.
(Continued next page.)

TABLE A1 (continued)
List of the Contributing Networks to ITACAext 2.0

Network Code	No. of Data	%	Name/Description	Abbreviated Reference*	DOI
6N	193	0.4	Emersito seismic network in the Ancona area (Italy)	EMERSITO WG (2024)	10.13127/sd/qctgd6c-3a
7C	8	<0.1	Urban seismic network (RUSN-CAL-INGV)	Costanzo *et al.* (2022)	10.13127/sd/o5nrhmp9n7
7F	46	0.1	Snow avalanche seismic monitoring around Sibilla Mount (central Italy) (Snowbilla)	Marzorati *et al.* (2021)	10.13127/sd/fcjdr8key3
8C	2	<0.1	Mont-Blanc (Mont-Blanc)	Helmstetter *et al.* (2020)	10.15778/RESIF.8C2019
AC	21	<0.1	Albanian seismological network (AC)	Institute of GeoSciences, Polytechnic University of Tirana (IGEO PUT)	10.7914/SN/AC
BA[†]	73	0.1	Università della Basilicata, Italy (UniBas)	–	–
BW	2	<0.1	BayernNetz (BW)	LMU Munich (2001)	10.7914/SN/BW
C4	3	<0.1	CERN seismic network	CERN (2016)	10.12686/sed/networks/c4
CH	1829	3.5	Switzerland Seismological Network	SED (1983)	10.12686/sed/networks/ch
E	82	0.2	Agenzia Nazionale per le Nuove Tecnologie (ENEA) seismic network	–	–
EU	86	0.2	Generic European strong motion network	–	–
FR	470	0.9	RESIF and other broadband and accelerometric permanent networks in metropolitan France (RESIF)	RLBP (1995)	10.15778/RESIF.FR
G	7	<0.1	French Global network of seismological broadband stations (GEOSCOPE)	IPGP (1982)	10.18715/GEOSCOPE.G
GE	2	<0.1	GEOFON	GEOFON (1993)	10.14470/TR560404
GU	513	1	Regional seismic network of northwestern Italy (RSNI)	UniGE (1967)	10.7914/SN/GU
HL	1	<0.1	National Observatory of Athens seismic network	NOA (1975)	10.7914/SN/HL
IT	17981	34.1	Italian strong motion network (RAN)	DPC (1972)	10.7914/SN/IT
IV	23664	44.9	Italian national seismic network (INSN)	Istituto Nazionale di Geofisica e Vulcanologia (INGV) (2005)	10.13127/sd/x0fxnh7qfy
IX	231	0.4	Irpinia seismic network (ISNet)	–	–
IY	6	<0.1	Rete Sismica Università della Calabria	Unical (1981)	10.7914/SN/IY
ME	12	<0.1	Montenegrin seismic network (ME)	Montenegrin Seismic Network (1982)	10.7914/SN/ME
MN	1239	2.4	Mediterranean very broadband seismographic network (MedNet)	MedNet (1990)	10.13127/sd/fbbbtdtd6q
NI	151	0.3	Northeast Italy broadband network (NI)	OGS (2002)	10.7914/SN/NI

*The full citation is reported in the reference list.
[†]The network code is now changed to 1V (Gallipoli, 2003). Full citation in the reference list.
(Continued next page.)

TABLE A1 (continued)
List of the Contributing Networks to ITACAext 2.0

Network Code	No. of Data	%	Name/Description	Abbreviated Reference*	DOI
OE	6	<0.1	Austrian seismic network	ZAMG (1987)	10.7914/SN/OE
OX	293	0.6	Northeast Italy seismic network	OGS (2016)	10.7914/SN/OX
RA	704	1.3	Réseau Accélérométrique Permanent (French Accelerometrique Network) (RAP)	RESIF (1995)	10.15778/RESIF.RA
RF	176	0.3	Friuli Venezia Giulia Accelerometric Network (RF)	University of Trieste (1993)	10.7914/SN/RF
RO	5	<0.1	Romanian Seismic Network (NIEP)	NIEP (1994)	10.7914/SN/RO
RR	3	<0.1	Reti accelerometriche centri storici italiani (RACSI)	RACSI (2017)	10.7914/SN/RR
SI	71	0.1	Province Südtirol	–	–
SL	234	0.4	Seismic network of the Republic of Slovenia (SNRS)	Slovenian Environment Agency (1990)	10.7914/SN/SL
ST	127	0.2	Trentino seismic network	Geological Survey-Provincia Autonoma di Trento (1981)	10.7914/SN/ST
TV	253	0.5	INGV experiments network	–	–
X3	5	<0.1	INGV SISMIKO emergency seismic network for Salemi, Italy (ISesnSI)	Alparone *et al.* (2020)	10.13127/sd/ljfpeosek0
XE	6	<0.1	Pollino broadband near-fault Observatory (2022–2033), southern Italy (Pollino BB-NFO)	Chiarabba *et al.* (2022)	10.13127/sd/ns6hb3q8ve
XK	19	<0.1	Temporary network seismic in central Calabria (TeCC)	–	–
XO	714	1.4	Seismic network for site-effect studies in Amatrice area (central Italy) (SESAA)	EMERSITO WG (2018)	10.13127/sd/7txegdo5x8
Y1	164	0.3	XVII-INGV SISMIKO emergency seismic network (ISESN) (XVII-ISESN)	D'Alema *et al.* (2022)	10.13127/sd/tblkba-3u6
YD (2003–2006)	25	0.1	Calabria-Appennine-Tyrrhenian/ Subduction-Collision-Accretion Network (CATSCAN)	Art Lerner-Lam (2003)	10.7914/SN/YD_2003
YD (2018)	2	<0.1	INGV SISMIKO emergency seismic network for Molise, Italy (ISESMI)	Moretti *et al.* (2018)	10.13127/sd/fir72chywu
YI	4	<0.1	Retreating-trench, extension, and accretion LTectonics: A multidisciplinary study of the northern Apennines (RETREAT)	Vadim (2003)	10.7914/SN/YI_2003
YP (2012–2013)	114	0.2	CIFALPS	Zhao *et al.* (2016)	10.15778/RESIF.YP2012
YP (2014–2015)	3	<0.1	Argostoli earthquake aftershock experiment	Cornou *et al.* (2014)	10.15778/RESIF.YP2014
Z3	2	<0.1	AlpArray	AlpArray Seismic Network (2015)	10.12686/alparray/z3_2015

*The full citation is reported in the reference list.
†The network code is now changed to 1V (Gallipoli, 2003). Full citation in the reference list.
(Continued next page.)

TABLE A1 (continued)
List of the Contributing Networks to ITACAext 2.0

Network Code	No. of Data	%	Name/Description	Abbreviated Reference*	DOI
ZH	6	<0.1	INGV SISMIKO Emergency Seismic Network for Chianti Fiorentino - Italy (ISESNCFI)	Piccinini *et al.* (2022)	10.13127/sd/jhggg3bnlq
ZM	7	<0.1	Seismic Emergency for Ischia by Sismiko (SEISk)	Galluzzo *et al.* (2017)	10.13127/sd/u0j8hsaevf
ZN	6	<0.1	SISMOVALP	–	–
ZO	14	<0.1	Dynamic Planet (PD) multiparametric Northern Italy seismic network (PDNET)	Massa *et al.* (2021)	10.13127/sd/yhcfomcbo_

*The full citation is reported in the reference list.
†The network code is now changed to 1V (Gallipoli, 2003). Full citation in the reference list.

Appendix B

Details on ITA18 functional form

The functional form of ITA18 (Lanzano, Luzi, *et al.*, 2019; Lanzano *et al.*, 2022) is

$$\log_{10} Y = a + F_M(M_w,\mathrm{SOF}) + F_D(M_w,R) + F_S(V_{S,30}) + \sigma, \quad \text{(B1)}$$

$$F_M(M_w,\mathrm{SOF}) = f_j \mathrm{SOF}_j + b_1(M_w - M_h)(M_w \le M_h) + b_2(M_w - M_h)(M_w > M_h), \quad \text{(B2)}$$

$$F_D(M_w,R) = [c_1(M_w - M_{\mathrm{ref}}) + c_2]\log_{10}\sqrt{R^2+h^2} + c_3\sqrt{R^2+h^2}, \quad \text{(B3)}$$

$$F_S(V_{S,30}) = k \cdot \log_{10}\frac{V_0}{800}, \quad \text{(B4)}$$

$$\sigma^2 = \tau^2 + \phi_{S2S}^2 + \phi_0^2, \quad \text{(B5)}$$

in which Y is the observed IM, that is, the peak ground acceleration and velocity (PGA and PGV) and 36 ordinates of acceleration response spectra at 5% damping (SA) in the period (T) range 0.01–10 s, or 81 ordinates of the Fourier spectrum (FAS) in the frequency range (f) 0.1–30 Hz. The prediction is valid for RotD50, which is the median of the distribution of the intensity measures (IMs), obtained from the combination of the two horizontal components across all nonredundant azimuths (Boore, 2010).

The explanatory variables are the moment magnitude M_w, the source-to-site distance R (for both the Joyner–Boore and rupture distance metric), the shear-wave velocity $V_{S,30}$ and the styles of faulting SOF_j, which are dummy variables, introduced to specify strike-slip (j = 1), reverse (j = 2), and normal (j = 3) fault types. In particular, $V_0 = V_{S,30}$ for $V_{S,30} < 1500$ m/s and $V_0 = 1500$ m/s for $V_{S,30} < 1500$ m/s.

The hinge magnitude M_h, the reference magnitude M_{ref}, and the pseudodepth h are obtained in a first step nonlinear regression. The coefficients a, b_1, b_2, c_1, c_2, c_3, k, and f_j (f_1 for strike-slip fault, f_2 for thrust fault, and f_3 for normal fault) are calibrated by a mixed-effect linear regression (Bates *et al.*, 2015). The random-effects are applied to stations and events, to estimate the partially nonergodic sigma according to Al Atik *et al* (2010), in which τ and φ_{S2S} represent between-event and site-to-site variability, respectively, and φ_0 is the standard deviation of the event- and site-corrected residuals. The coefficients were provided in Lanzano, Luzi, *et al.* (2019) and Lanzano *et al.* (2022).

Manuscript received 7 August 2024
Published online 9 June 2025