



# Adjusting an active shallow crustal ground motion model to regions with scarce data: application to France

F. Ramadan<sup>1,2</sup> · G. Lanzano<sup>1</sup> · F. Pacor<sup>1</sup> · C. Felicetta<sup>1</sup> · C. Smerzini<sup>2</sup> · P. Traversa<sup>3</sup>

Received: 9 November 2023 / Accepted: 1 March 2024  
© The Author(s) 2024

## Abstract

The objective of this work is to test whether an empirical Ground Motion Model (GMM) developed for high-seismicity regions can be effectively adapted to a neighbouring region with lower seismic activity. We select the ITA18 suite of GMMs (Lanzano et al. in Bull Seismol Soc Am 109(2): 525-540, 2019a), developed for Italy, which is a region dominated by moderate-to-strong shallow crustal earthquakes, and assess their applicability to Continental France, where the seismic activity is less frequent and characterised by lower magnitudes. Based on a dataset of more than 2300 records of events with  $3.0 \leq M_w \leq 5.2$ , occurred in France in the time interval 1996–2019 (named FR20), we perform a residual analysis and calibrate an adjustment factor for both horizontal and vertical-to-horizontal (VH) components of Peak Ground Acceleration, Peak Ground Velocity, and 5% damped Spectral Acceleration (SA). Apart from the median correction, no modification of the scaling with magnitude, focal mechanism, and  $V_{S,30}$  is introduced, while the distance scaling is adjusted to capture the lower anelastic attenuation of the French data. In addition, to overcome the underestimation of the ITA18 model for the short period VH spectral amplitudes in the near-source region ( $R_{\text{epi}} < 15$  km), an additional empirical corrective factor is introduced. In spite of the good agreement of the adjusted model with respect to the median trends of the FR20 dataset, a regionalization of the source effects is introduced to reduce the relatively high between-event variability of the proposed model. The proposed model provides predictions similar to ITA18 in the most seismically active regions (Alps or Pyrenees), while, in the other zones, the predicted amplitudes are richer at high frequencies. Given the paucity of seismic records in these zones, this behavior should be confirmed on the basis of additional data (e.g. physics-based simulations, geologic and tectonic features). The use of the proposed model for hazard applications is recommended within the validity limits of the data ( $3.0 \leq M_w \leq 5.2$ ). However, the similarity of the ground motion in the Alps and Pyrenees with the predictions of ITA18 suggests that the adjusted model could be also employed for higher magnitudes, upon suitable checks.

---

✉ F. Ramadan  
fadel.ramadan@ingv.it

<sup>1</sup> Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Milano, Milan, Italy

<sup>2</sup> Department of Civil and Environmental Engineering, Politecnico di Milano, Milan, Italy

<sup>3</sup> Electricité de France (EDF – DIPNN – DT), Aix-en-Provence, France

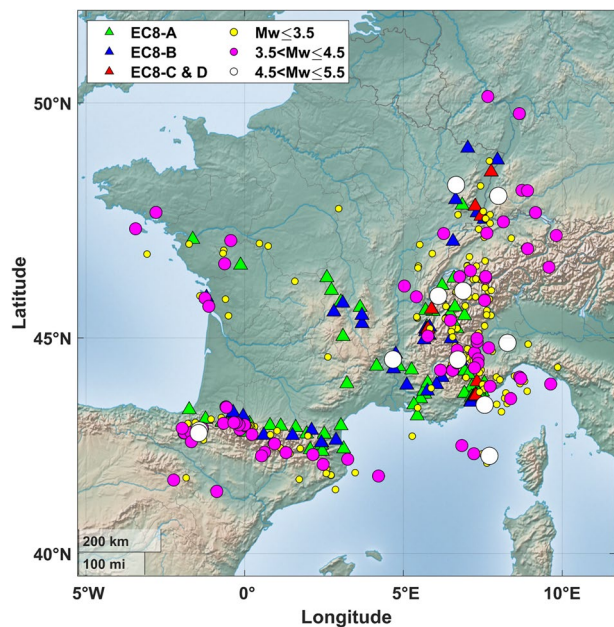
**Keywords** France · Ground motion model · Residual analysis · Strong motion dataset

## 1 Introduction

Continental France is a territory where seismic activity is sparse and characterised by weak-to-moderate events. Duverger et al. (2021) provides a comprehensive overview of instrumental seismicity in metropolitan France from 2010 to 2019. During this period, around 430 natural events ( $M \geq 3.0$ ) were recorded within the French territory, with only 4 events with  $M \geq 5.0$ . The seismicity is clustered in specific areas, such as the Alps (France, Italy and Switzerland borders), the Pyrenees (France and Spain border) and the Armorican Massif (North-Western region), while it is very rare in Central France and in the Paris and Aquitaine basins (Fig. 1).

On the other hand, Italy is characterised by medium to high seismic activity: in the same observation period (from 2010 to 2019), almost 4,000 events have been recorded, of which about 40 with local magnitude  $M \geq 5.0$  (<http://terremoti.ingv.it/>). The earthquake data recorded in Italy enabled the construction of robust ground motion models (GMMs) to support seismic hazard assessments as early as the late 1980s', starting with the prediction equations by Sabetta and Pugliese (1987) till the current model, ITA18, proposed by Lanzano et al (2019a). On the contrary, in France, due to the scarcity of data, different strategies have been tried to select or determine GMMs to be applied for hazard analyses. The most common practice consisted of selecting and using models developed in other regions of the world or at global scale which, however, might not be consistent with the ground motion of the target region, especially with regard to the assessment of associated variability (Cotton et al. 2006). More recently, however, few attempts to calibrate regional models for France have been done, due to the difficulty of dealing with the lack of observations. Drouet and Cotton (2015) used stochastic simulations based on

**Fig. 1** Map of events (circles), coloured by moment magnitude, and stations (triangles), coloured by EC8 soil category (CEN, 2004), of the FR20 dataset



a seismological model for France to develop a ground motion model for small to large magnitude events. Ameri et al. (2017) proposed a GMM for Europe, using the RESORCE dataset (Akkar et al. 2014) along with data from regions with small-to-moderate seismicity such as France and Switzerland, and accounting for the stress-parameter scaling in the ground motion modelling.

More recently, the development of partially or fully non-ergodic models has introduced the possibility to determine adjustment factors for existing GMM to account for differences in attenuation properties, source mechanisms and site conditions of the target region. The European-scale non-ergodic GMM by Kotha et al. (2020) proposed corrective terms for source and propagation effects, based on the Engineering Strong-Motion (ESM) data (Lanzano et al. 2019b). However, being the minimum magnitude of ESM dataset set to 4, the French data are poorly sampled by this dataset and, hence, this GMM is not able to capture all the differences across the French territory. Recently, a Bayesian update of the Kotha et al. (2020) model was performed using French records (Kotha and Traversa 2024). Keeping unchanged the regionalization used in the Kotha et al (2020) model, adjustment terms for source and propagation effects were computed for the French territory, which extended the applicability of the model to earthquakes recorded in France. Sung et al. (2023) proposed a fully non-ergodic GMM for the ordinates of the Fourier spectrum, using the French dataset compiled by Traversa et al. (2020), composed by velocimetric and accelerometric signals. Starting from the ergodic GMM of Bayless and Abrahamson (2019) for California, as backbone model, the Authors introduced a spatially-variable corrective model for the anisotropic path effects.

In line with these recent researches, the aim of this work is to assess ground motion for seismic hazard assessment in areas like France, i.e., with low seismic activity or, in general, regions where recorded data could be insufficient to determine robust empirical GMMs. In this work, we apply the “Referenced Empirical approach” introduced by Atkinson (2008) to make the predictions of a robust GMM, such as those available for California, compatible with the observed ground motion for an area of low seismicity with few available data (Northern-Eastern America). Specifically, we evaluate the differences between the seismic ground motion in France and Italy by means of the residual analysis of the ground motion dataset prepared by Traversa et al. (2020) with respect to the reference GMM for shallow active crustal events in Italy (ITA18, Lanzano et al. 2019a). The analysis is carried out for peak ground acceleration (PGA), peak ground velocity (PGV), and several ordinates of 5% damped acceleration response spectra (SA), both for horizontal and vertical components of the ground motion. Analysis of the results makes it possible to build a functional form and calibrate a model that adjusts ITA18 predictions, with the goal of making them usable in continental France for hazard studies and engineering analysis.

## 2 FR20 dataset

The FR20 dataset was built starting from the dataset published by Traversa et al. (2020), which includes records from broadband and accelerometric sensors provided by permanent and temporary seismic networks operated by French research institutions and partners grouped within the Réseau Sismologique et géodésique Français (RéSif) consortium (RESIF 1995) from 1996 up to the end of 2019.

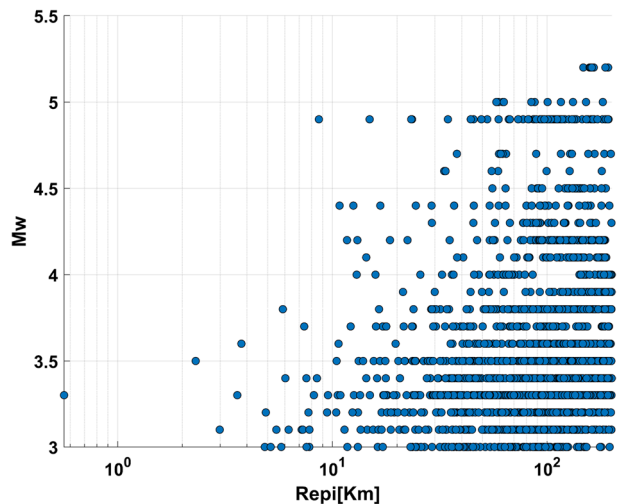
We selected events with moment magnitude ( $M_w$ ) greater than or equal to 3.0 and epicentral distance ( $R_{\text{epi}}$ ) shorter than 200 km; we mainly considered waveforms recorded

by accelerometric sensors (channel HN), and we included data from a few verified velocimetric sensors. We selected only data with 3 ground-motion components. Moreover, we selected stations installed in free-field or free-field-like conditions, in order to minimize the soil-structure interaction effects. Regarding the depth of the sensor, we consider stations installed down to 3 m below the ground surface. For the acceleration response spectra, 20 spectral ordinates were selected in the period range  $T=0-10$  s. The final dataset includes 2327 records from 290 earthquakes recorded by 121 stations during the period 1996–2019.

Figure 1 shows the spatial distribution of the events and the stations of the FR20 dataset: the majority of the events occurred close to the borders with Italy, Switzerland and Spain. The  $M_w$  ranges between 3.0 and 5.2 (Fig. 2), with most records in the range 3.0–4.0 and the focal depth varies between 0 and 30 km. About 50 waveforms are recorded at  $R_{epi}$  lower than 15 km, where we could expect to observe near-source effects. Among the events with the largest magnitudes in the dataset, the  $M_w$  4.9 earthquake of November 11, 2019, which occurred close to Montélimar city within the Rhone river valley, presents a considerable number of records. In the last years, this event has attracted attention due to its relatively large magnitude and the proximity to two critical facilities, namely the Cruas Nuclear Power Plant (NPP) at  $R_{epi}$  of about 15 km and Tricastin NPP at  $R_{epi} = 25$  km. This event is characterised by a reverse fault mechanism and it is very shallow (Ritz et al. 2020; Causse et al. 2021; Cornou et al. 2021), with hypocentral depth of about 1 km. It was recorded by about 30 stations within 200 km and only four records are within  $R_{epi} < 50$  km, namely, CRU1, BANN, OGLP and TRI2. The closest station, CRU1 ( $V_{S,30} = 660$  m/s), located at about 15 km from the epicentre, has recorded the largest PGA and PGV, equal to  $44 \text{ cm/s}^2$  and  $1.3 \text{ cm/s}$ , respectively.

The original dataset does not provide the Style of Faulting (SOF) of the events, but the reference GMM requires SOF as input variable. In order to infer such parameter for FR20 events, we used the seismotectonic zonation implemented in the probabilistic seismic hazard map for metropolitan France by Drouet et al. (2020), which provides, for each zone, the most likely type of focal mechanism (Strike-Slip, Normal, Reverse). In case of mechanisms with equal probabilities, the event is assumed to be strike-slip, as this is the most frequent style of faulting in France (Mazzotti et al. 2021; <https://data.oreme.org/observation/>

**Fig. 2** Moment Magnitude ( $M_w$ )—epicentral distance ( $R_{epi}$ ) distribution of the FR20 dataset



fmhex). As a result, around 52% of the events (150) are classified as strike-slip, 30% (87) are reverse and the remaining 18% of the events (53) have a normal faulting mechanism (Fig. 3a).

Concerning the site response characterization of the recording sites, a value for the average shear wave velocity in the uppermost 30 m ( $V_{S,30}$ ) is associated to all the selected recording stations. Among the 121 selected stations, 24 (about 20%) have the  $V_{S,30}$  values directly obtained from measured  $V_S$  profiles beneath the site, as provided by Traversa et al (2020), see Fig. 3b. The remaining  $V_{S,30}$  estimates are inferred from different proxies, i.e., the surface geology (57%) and topographic slope (23%).

### 3 ITA18 ground motion model

During the last years, efforts have been made in Italy to update ground motion models for crustal active areas, starting from the GMM of Bindi et al. (2011), calibrated soon after the 2009 L'Aquila seismic sequence. The initiative started in 2018, after the 2016–2017 sequence in central Italy, which provided a great amount of data from medium to large earthquakes and recordings at near-source distances (Luzi et al. 2017). The related research activity led to the compilation of a dataset for calibrating GMMs (Lanzano et al. 2022), which includes 5778 waveforms, related to 156 events in the magnitude range 3.5–8 and 1684 recording stations located within 200 km from the epicentre. All the stations are characterized by a value of  $V_{S,30}$ , that can be preferably measured or, alternatively, inferred from topography and geological maps, with similar criteria adopted for the French stations. The resulting  $V_{S,30}$  mainly corresponds to EC8 (CEN 2004) soil class B, i.e.,  $V_{S,30}$  included in the interval 360–800 m/s.

The dataset was first used by Lanzano et al. (2019a) to calibrate a partially non-ergodic predictive model of PGA, PGV, and 36 ordinates of SA in the 0.01–10 s interval, here named ITA18-H, for both the Joyner-Boore ( $R_{JB}$ ) and rupture ( $R_{rup}$ ) distances. More recently, Ramadan et al. (2021) calibrated a GMM to predict the vertical-to-horizontal ratio for the same intensity measures (ITA18-VH). Then, the vertical component of the ground motion (ITA18-V) can be obtained from the multiplication of the predictions of ITA18-H by the corresponding ITA18-VH

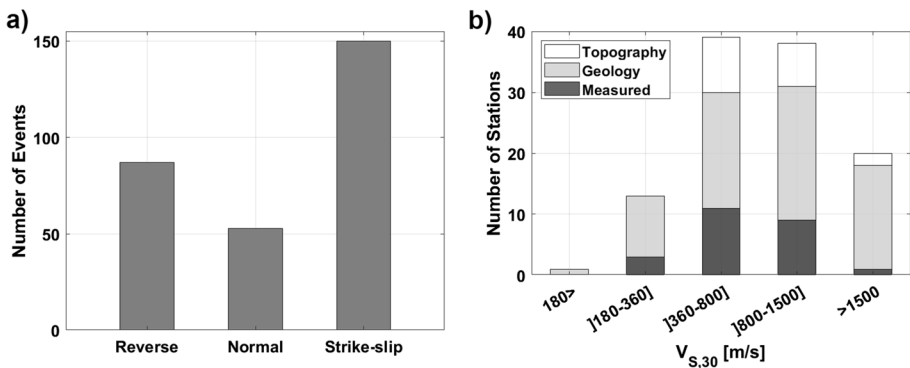


Fig. 3 FR20 data distributions. **a** earthquake focal mechanisms; **b** station  $V_{S,30}$  values according to the type of estimates

predictions for the same scenario. For sake of completeness, a brief description of the functional forms of ITA18-H and ITA18-VH models is provided in the Appendix 1.

## 4 Residual analysis of FR20 dataset with respect to ITA18 GMMs

### 4.1 Method

The total residuals between the ITA18 predictions and FR20 data are calculated according to the following expression:

$$R_{es} = \log_{10} Y_{\text{Obs-FR20}} - \log_{10} Y_{\text{Pred-ITA18}} \quad (1)$$

where,  $R_{es}$  are the total residuals obtained as the base 10 logarithm difference between FR20 observations ( $Y_{\text{Obs-FR20}}$ ) and predictions of ITA18 ( $Y_{\text{Pred-ITA18}}$ ). The variable  $Y_{\text{Obs-FR20}}$  refers to the different intensity measures (PGA, PGV and SA) of the FR20 dataset. For the horizontal component, the geometric mean of NS and EW components is considered.

Because the data processing is manual (see Traversa et al. 2020 for details on the data processing), FR20 high- and low-pass filter corner frequencies may vary. As a result, the number of records usable for residual analysis varies with periods (see Figure A1 in the electronic supplement #1): in particular, the number of usable records starts to decrease at periods higher than 1 s and reduces by about 40% (about 900 records) at  $T=2$  s. As the reduction of recordings is significant at long periods, in the following we will limit our analysis to a range of periods  $T=0.01-2$  s for SA.

Since the geometries of the faults are not available in the FR20 dataset and considering that for small earthquakes we can assume a point-like source, the epicentral distance  $R_{\text{epi}}$  is used in place of  $R_{\text{JB}}$  and the hypocentral distance ( $R_{\text{hyp}}$ ) in place of  $R_{\text{rup}}$  for comparisons with ITA18 predictions.

In order to investigate the dependence of the residuals with respect to the explanatory variables, the total residuals are decomposed according to the partially non-ergodic approach for GMMs calibration (Al Atik et al. 2010; Rodriguez- Marek et al. 2014). In particular:

$$R_{es} = \delta_0 + \delta B_e + \delta S2S_s + \delta W_{es} \quad (2)$$

where total residuals are separated into median bias ( $\delta_0$ ), between-event ( $\delta B_e$ ), site-to-site ( $\delta S2S_s$ ) and site- and event- corrected ( $\delta W_{es}$ ) terms and the subscripts denote a systematic effect for event  $e$  and/or station  $s$ . The  $\delta B_e$  represents the average misfit of the observed ground motion of an individual earthquake, from the median population predicted by the GMM. The  $\delta S2S_s$  represents the systematic deviation of the observed ground motion at a station (i.e., the site term) from the median prediction of the GMM. Finally,  $\delta W_{es}$  is the site- and event- corrected residual, i.e. the leftover residual after systematic event and site residuals are removed. The residual components are evaluated by means of the mixed-effect modelling (Stafford 2014; Bates et al. 2015), as is commonly done for the calibration of non-ergodic GMMs. The analyses are conducted with the *fitlme* function, made available in Matlab ® computational software, which allows performing linear mixed-effects regressions.

### 4.2 Results

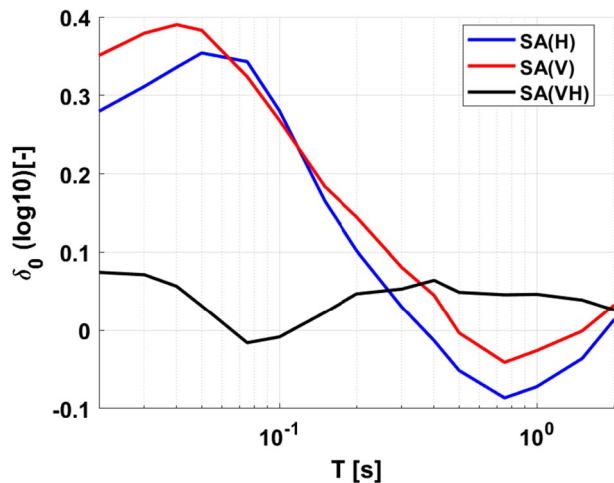
Figure 4 shows the graph of the FR20 total bias (offset),  $\delta_0$  in Eq. (2), for SA (H, V and VH components) as a function of period T. The reference model is the ITA18 GMM in Joyner-Boore distance  $R_{JB}$ , but the findings are similar for rupture distance. Positive offset indicates that the reference model underestimates, on average, the observations, while negative values the opposite.

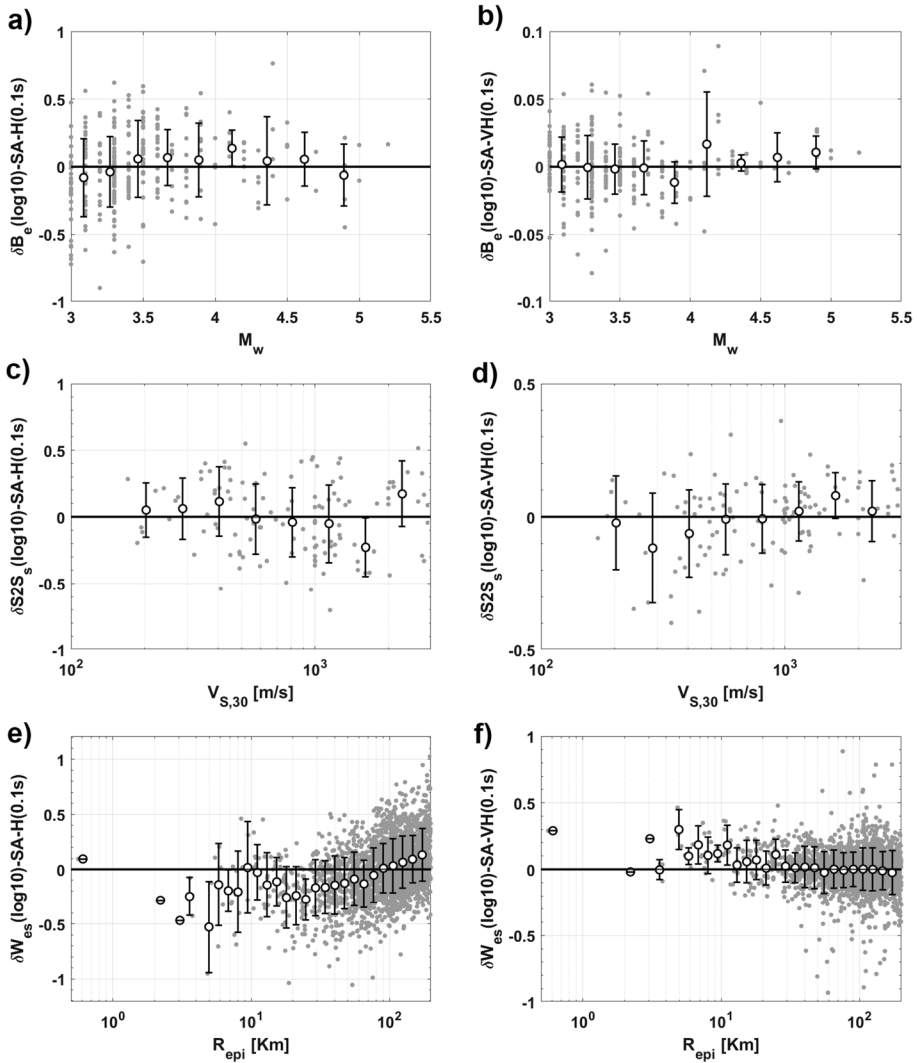
Since  $\delta_0$  is found to be positive, ITA18-H predictions on average underestimate FR20 observations at short periods ( $T < 0.2$  s), both for horizontal and vertical components of SA. Conversely, at long periods, a slight overestimation is observed (negative  $\delta_0$ ). Considering the similar trends of  $\delta_0$  for the vertical and horizontal components, the overall bias associated with the ratio VH is close to zero or slightly positive, which indicates a small underestimation by ITA18-VH predictions. The observed differences between the median seismic motion in France and Italy seem in agreement with the GMM of Kotha et al. (2020) and Kotha and Traversa (2024), which introduced a very broad regionalization of seismic sources and seismic attenuation properties at European scale: at short periods, the seismic motion in France is systematically higher than that observed in Italy, especially with respect to the Central Apennines, whose seismicity contributed crucially to the calibration of ITA18.

Aside from median offset  $\delta_0$ , which is based on all events, distances and site conditions, the other residual terms in Eq. (2), namely  $\delta B_e$ ,  $\delta S2S_s$  and  $\delta W_{es}$ , can be separately examined to test whether and to what extent the ITA18 models are able to capture the magnitude, distance and  $V_{S,30}$  scaling of the FR20 dataset. Figure 5 shows the trends of  $\delta B_e$  with magnitude,  $\delta S2S_s$  with  $V_{S,30}$  and  $\delta W_{es}$  with the epicentral distance for H and VH components of SA at short periods ( $T=0.1$  s). In the electronic supplement #1 (see Figure A2), the same plots of Fig. 5 for SA-T=1 s are also provided. In addition, graphs of the residuals for the vertical component (SA-T=0.1 s and 1 s) are also presented in the electronic supplement #1 (see Figure A3), to show that the trends are similar to those observed for the horizontal component.

The between-event components  $\delta B_e$  do not show any significant bias with  $M_w$ , regardless of the period under consideration, indicating that the magnitude scaling of FR20 data is

**Fig. 4** Median offset of the FR20 residuals (Eq. 2) for horizontal, vertical and VH ratio components of SA ordinates





**Fig. 5** FR20 horizontal (a, c, and e) and VH (b, d, and f) residuals for SA-T=0.1 s vs different explanatory variables: a and b  $\delta B_e$  vs. moment magnitude; c and d  $\delta S2S_s$  vs  $V_{s,30}$ ; e and f  $\delta W_{es}$  vs  $R_{epi}$

fairly well captured by the ITA18 model, both for horizontal and vertical components (Figs. 5a and d), even for magnitudes lower than 3.5 which are scarcely represented in the ITA18 dataset. The  $\delta B_e$  values are affected by a large variability, which is likely related to the uncertainties and regional variabilities associated to the source parameters provided of French earthquakes (see also discussion in Kotha & Traversa 2024).

The  $\delta S2S_s$  values are affected by a large variability (Figs. 5b and 5e), which is probably due to the fact that  $V_{s,30}$  for many sites is inferred using proxies. However, the  $V_{s,30}$  scaling of the ITA18 models seems adequate to describe the site effects in France, since the average values of  $\delta S2S_s$  are not affected by significant bias at both short and long periods.



The only relevant biases concern the underestimation observed at  $V_{S,30} \geq 1500$  m/s at short period, that can be related to site scaling of the reference GMM: indeed, the scaling with  $V_{S,30}$  is linear up to the threshold value,  $V_{S,30} = 1500$  m/s, beyond which the site term becomes constant. On one hand, this kind of site, representative of hard-rock conditions, is scarcely present in Italy, where the rocks have less rigid mechanical properties (Forte et al. 2019; Mori et al. 2020). On the other hand, such bias could also be related to the fact that only 3 stations with  $V_{S,30} \geq 1500$  m/s dispose of measured  $V_{S,30}$ , while the others have values inferred from other proxies. For long periods (see electronic appendices), the  $\delta S2S_s$  of horizontal components show positive values, indicating general overestimation of ground motion for sites with low  $V_{S,30}$ .

Regarding the  $\delta W_{es}$  median trend with distance (Figs. 5e and 5f), the ITA18-H model tends to attenuate faster with respect to FR20 short-period data at distances larger than 100 km (positive residuals). The opposite occurs at shorter distances, where the event- and site- corrected residuals are on average negative. Although limited to few zones within south-eastern France and the Pyrenees, the faster average ground motion attenuation in Italy with respect to France is also observed by Kotha and Traversa (2024), particularly at short periods. Using an independent attenuation tomography approach based on coda waves, Mayor et al. (2018) also suggested higher attenuation in Italy than in France at long distances.

The VH event- and site- corrected residuals show a much lower bias (and associated variability) than the horizontal residuals: in epicentral area (less than 15 km) the residuals are positive, which suggests a possible underestimation of the vertical ground motion by ITA18 in the near field. This was also observed in the calibration of ITA18-VH (Ramadan et al. 2021) with respect to the NESS dataset (Pacor et al. 2018), i.e., a dataset of near-source waveforms of global earthquakes of  $M_w \geq 5.5$ . To overcome this prediction bias, Ramadan et al. (2021) calibrated a correction coefficient to adjust the near-field predictions (the corrected model is named ITA18-VH-NESS).

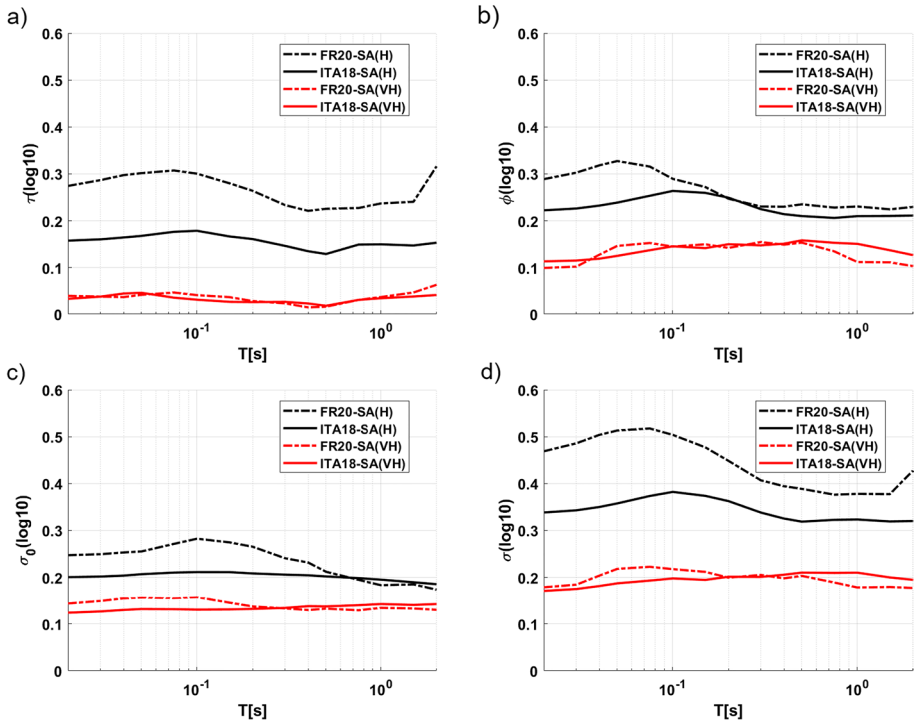
The variability of the French ground motion is computed from the standard deviations of the residual terms in Eq. (2). In particular, the total (ergodic) standard deviation is calculated according to the following equation:

$$\sigma = \sqrt{\tau^2 + \phi^2 + \sigma_0^2} \quad (3)$$

where  $\tau$ ,  $\phi$  and  $\sigma_0$  are the standard deviations associated with  $\delta B_{es}$ ,  $\delta S2S_s$  and  $\delta W_{es}$ , respectively.

Figure 6 shows the standard deviations in Eq. (3) (total and components) for SA ordinates, compared with ITA18, for both horizontal and VH components. The total standard deviation from the French data is significantly larger than those of the ITA18 models. This difference is mainly related to the between-event component  $\tau$ , which is almost twice as those of ITA18. Such large variability of the event term was already observed by other authors (Ameri et al. 2017, Sung et al. 2023, Kotha and Traversa 2024) and could be related to multiple causes:

- a) The different seismotectonic features of France with respect to Italy, which includes areas with frequent active crustal seismicity (Alps, Pyrenees) and areas with more stable continental seismicity in the north-west of the country. Moreover, the calibration datasets FR20 and ITA18 have different  $M_w$  ranges, since FR20 is shifted toward small



**Fig. 6** **a** Between-event ( $\tau$ ), **b** site-to-site ( $\phi$ ), **c** event- and site-corrected ( $\sigma_0$ ) and **d** total ( $\sigma$ ) standard deviations for H and VH FR20 residuals compared to the ITA18 sigma's

- and intermediate values (3.0–5.2) when compared to the range of the ITA18 dataset (3.5–7.5);
- b) Magnitude estimates of FR20 are affected by a significant uncertainty, related to the fact that  $M_w$  estimates are not homogeneously estimated over the FR20 earthquakes (see Traversa et al. 2020 for details). As a matter of fact, previous works suggested possible lack of homogeneity in the  $M_w$  estimates of the Si-Hex catalogue (Cara et al. 2015), associated to the FR20 earthquakes, with respect to other estimates (see for example Ameri et al. 2017). Cara et al. (2015) also highlighted a lack of homogeneity in the  $M_w$  estimates for different magnitude ranges within the Si-Hex catalogue itself. Earthquake depths are also poorly constrained in France, mainly due to a less dense station coverage than Italy and to the fact that earthquakes are mostly located close to the borders and the coasts;
  - c) Finally, although scaling with magnitude showed no trend, the large event variability suggests a role of other source parameters, such as stress drop, focal depth, etc. Indeed, the regional or local difference in terms of stress drop is indicated by several authors (e.g. Drouet et al 2010; Ameri et al. 2017) as one of the causes of the highest variability of the event term at high frequencies.

Regarding the variability of the site-to-site term,  $\phi$ , the one observed for the FR20 dataset is higher than ITA18 only at short periods, probably due to the quite large variability among the French sites (see Fig. 5c), due to the greater presence of hard-rock sites

in France, with less attenuating properties than the rock in Italy (Cotton et al. 2006; Drouet et al. 2010). The variability obtained from the aleatory residual of FR20, i.e.  $\sigma_0$ , is also higher than that obtained from ITA18, arguably due to the differences between the average seismic motion attenuation in France and Italy, observed in Fig. 5e.

The values of the standard deviations of the components of the residuals for VH are very similar to those obtained by Ramadan et al. (2021) for ITA18, confirming that the VH variabilities are always smaller than those obtained for horizontal component and much less influenced by the regional characteristics of the seismic motion.

## 5 Adjustment models for France

Based on the previous findings, period-dependent correction factors are calibrated to adjust ITA18 predictions to the FR20 dataset, according to these criteria:

- The prediction for horizontal components takes into account (a) a global corrective term (bias), (b) a correction for the scaling with distance and (c) a set of corrective coefficients to capture the ground motion peculiarities of the different seismogenic areas of France and reduce the event variability;
- The prediction for VH ratio is only adjusted for short-distance conditions ( $R < 15$  km).

### 5.1 Model for the horizontal components

Based on the trend of residuals observed in the previous section for the horizontal component of seismic motion in France, an adjustment term  $\delta_H$  is calibrated, such that the predictions for the French territory can be obtained using the model:

$$\log_{10} Y_{\text{IFR20-H}} = \log_{10} Y_{\text{ITA18-H}} + \delta_H \tag{4}$$

The adjustment term  $\delta_H$  is described by the following equation:

$$\delta_H = \delta_a + \delta_{c3} \sqrt{R^2 + h^2} + \delta L2L_z|_{\text{source zones}} + \delta B_e|_{\text{events}} + \delta S2S_s|_{\text{stations}} + \delta W_{es} \tag{5}$$

where the period-dependent coefficients are:  $\delta_a$  is the global bias with respect to ITA18 (Fig. 4);  $\delta_{c3}$  is a fixed-effect correction of the  $c_3$  coefficient of ITA18 to account for anelastic attenuation differences (see Appendix 1);  $\delta L2L_z$  is the so-called location-to-location random-effect term (e.g., Kotha et al. 2020), i.e., the systematic bias of the ground motion, observed for events originating at a given source (or source zone) with respect to the median prediction of a GMM. For the model proposed in Eq. (5), each  $\delta L2L_z$  refers to a source zone  $z$ , defined by the seismogenic zonation proposed in the next section (Sect. 5.1.2).  $R$  and  $h$  represent the source-to-site distance and pseudo-depth, respectively. Note that the values of  $h$  are not recalibrated but assumed to be the same as those proposed for ITA18 for the sake of simplicity.  $\delta B_e$ ,  $\delta S2S_s$ , and  $\delta W_{es}$ , already been introduced in the previous section, are recalculated considering the functional form in Eq. (5). Two sets of coefficients are provided in electronic supplement #2, by considering  $R$  as epicentral ( $R_{\text{epi}}$ ) and/or hypocentral ( $R_{\text{hyp}}$ ) distances, respectively.

### 5.1.1 Anelastic attenuation correction

To adjust ITA18-H predictions by considering the differences in attenuation characteristics observed for the French data, a corrective term for the anelastic attenuation coefficient is estimated through regression on all FR20 data. Figure 7 shows the trend of the anelastic attenuation coefficient  $c_3$  (see functional form in the Appendix 1) for Italy and France. In the first case, we refer to the  $c_3$  obtained from the ITA18 calibration, while for the French territory, the global attenuation is obtained by summing the  $c_3$  of ITA18 with the  $\delta_{c_3}$  obtained from the calibration of the corrective model in Eq. (5).

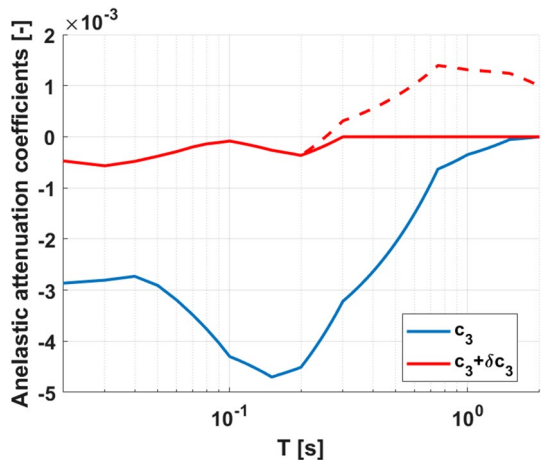
The global attenuation coefficient is larger for the French dataset with respect to the ITA18 one over the whole period range. From a physical point of view, this means that the median attenuation for the French territory is always slower than that observed for Italy. However, it should be noted that, for periods larger than 0.25 s,  $c_3 + \delta_{c_3}$  for France becomes weakly positive (dotted line in figure): this is an undesirable effect in GMM calibration that is quite commonly found at long periods. Specifically, when the attenuation decays with an upward rather than a downward concavity, at very long distances, it may occur that the ground motion increases at very long distances rather than decreases. Lanzano et al. (2019a) reports the occurrence of this undesirable effect also in the ITA18 calibration, which can be resolved by imposing that the anelastic attenuation term is zero, when it becomes positive. In the calibration of the model in Eq. (5) we then impose that  $c_3 + \delta_{c_3} = 0$  for  $T > 0.3$  s and re-adjust the other coefficients (fixed and random terms).

### 5.1.2 Source zonation

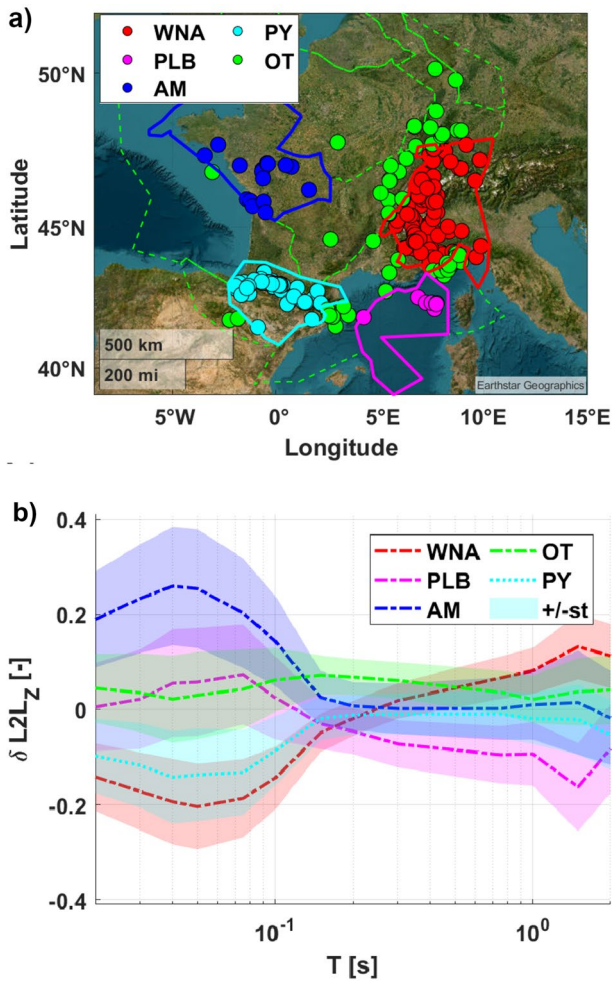
Recently, a seismogenic zonation for French territory was introduced by Drouet et al (2020) for seismic hazard assessment. The zonation proposes to aggregate the seismic source zones for the French territory into 11 domains, based on geological, seismotectonic, seismological and geophysical information (see Fig. 8a).

The map in Fig. 8a shows the FR20 earthquake epicenters which are located in 8 out of the 11 domains, but some of them are sampled by very few events. As a first stage, a preliminary analysis is performed with the aim of merging some seismogenic domains based on the trend of the between-event residuals ( $\delta B_e$ ), following an approach similar

**Fig. 7** Anelastic attenuation coefficient of ITA18 ( $c_3$ ) and the one adjusted for France ( $c_3 + \delta_{c_3}$ ) as a function of period (dotted line is the original regression; continuous line, the readjusted one)



**Fig. 8 a** Seismotectonic map proposed by Drouet et al. (2020) and FR20 event epicentral locations. **b** location-to-location random effects term  $\pm$  standard deviation as a function of period for each domain (colours correspond to those in Fig. 7a)



to Brunelli et al. (2023) for Italy. The  $\delta B_e$  estimates, computed using Eq. (2) without considering the random effect on the source zonation, are first averaged per domain and then merged on the basis of the similar average trend. The groupings are also supported by geological and structural considerations (e.g. the similarity of focal mechanisms). In this way, the number of  $\delta L2L_z$  effects is reduced, and we finally consider 5 source zones in Eq. (5) (Fig. 8a):

1. Western and Northern Alps (WNA);
2. Provence—Ligurian basin (PLB);
3. Armorican Massif (AM);
4. Pyrenees (PY);
5. Others (OT), grouping all the events that are not located in the previous zones.

Figure 8b shows the  $\delta L2L_z$  values for the five source zones to emphasise the role played by the source zonation into the model.  $\delta L2L_z$  is positive at short periods for the Armorican

Massif, indicating that the ground motion is up to 1.8 times higher ( $\delta L2L_z=0.25$ ) with respect to the median model without correction coefficients for spectral ordinates at  $T=0.07$  s. The corrections for the Alps and Pyrenees have mirrored trends, with negative  $\delta L2L_z$  values, resulting in a maximum reduction of about 30%, at short periods, with respect to the median ITA18 model corrected for the global bias and regional attenuation. For Provence and the Ligurian basin, the bias is most important for the intermediate periods with maximum reduction of 43% for the spectral ordinate in the interval  $T=0.8-1.5$  s. Finally, the OT correction is on average null and thus, for these areas, the median model can be used without correction.

In order to verify the robustness of the  $\delta L2L_z$  estimates, we have also reported the spectral values of the Standard Error (SE) associated with the correction estimates for the different zones in Fig. 8. The uncertainty associated with  $\delta L2L_z$  varies with period but is higher at short ( $T < 0.1$  s) and long periods ( $T > 1$  s), while it is lower at intermediate periods. The SE values are not negligible, especially for some areas such as AM and PLB, but the maximum values do not exceed half the  $\delta L2L_z$ .

### 5.1.3 Model variability

The total standard deviation of I.FR20-H can be calculated as:

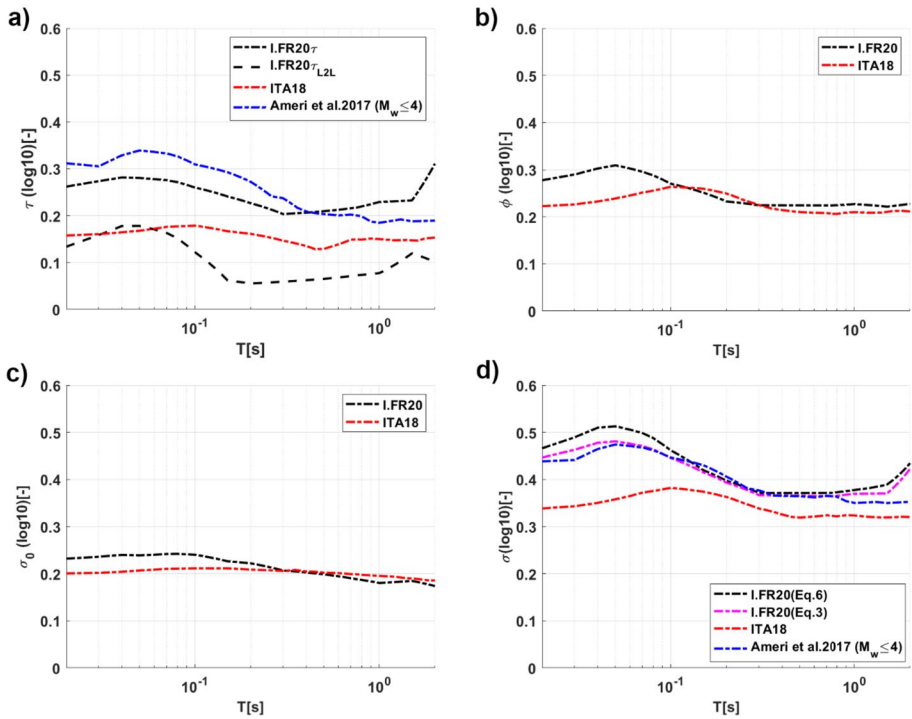
$$\sigma = \sqrt{\tau^2 + \phi^2 + \tau_{L2L}^2 + \sigma_0^2} \quad (6)$$

where the standard deviation of the  $\delta L2L_z$  term, called  $\tau_{L2L}$ , is also included, while the other terms refer to the standard deviations of the event term ( $\tau$ ), site ( $\phi$ ) and the leftover aleatory variability ( $\sigma_0$ ), defined in the previous section. On the other hand, to calculate the single-zone standard deviation,  $\sigma$ , i.e., the one associated with the prediction including the zone correction, we can refer to Eq. (3).

Figure 9 shows the different components of variability compared with those provided by ITA18 and the GMM by Ameri et al. (2017), considering their base-model (without stress-drop correction) with magnitude-dependent  $\tau$  (see Ameri et al. 2017 for details).

The standard deviation of  $\tau$  is reduced by 10% compared to that obtained from the analysis of the residuals in Fig. 6, and this result is due to the introduction of the zone correction, whose variability ( $\tau_{L2L}$ ) has non-negligible values, especially at short periods. However, the reduction of  $\tau$  is still small, and the reason may be twofold: the characteristics of the FR20 dataset still do not allow to significantly improve the performance of the model, because of the limited coverage of events, sites, and paths. Moreover, the differences between ITA18 and the model proposed here could be related to the quality of the metadata of the FR20 dataset, in particular the values of moment magnitude and event depth, as mentioned above. Furthermore, several non-ergodic models in the literature (Kotha et al. 2020; Sgobba et al. 2021) have shown that the introduction of the zone-based correction, although leading to an improvement in median estimates, often causes only a limited reduction of the event variability, which must be sought in other causes (e.g. stress-drop variability or poor quality of the event metadata). When compared to other models,  $\tau$  still remains higher than ITA18 at the short periods; on the contrary, the event standard deviation is lower than the heteroscedastic model by Ameri et al. (2017) for  $T < 0.7$  s, computed for  $M \leq 4.0$ .

The site-to-site standard deviation remains unchanged from the previous residual analysis, as we have not added a term in the correction  $\delta_H$  (Eq. [5]) to account for



**Fig. 9** Comparison between the standard deviations of the model proposed in this study for horizontal components and those by Ameri et al. (2017) and ITA18. **a** between-event ( $\tau$ ) and location-to-location ( $\tau_{L2L}$ ), **b** site-to-site ( $\phi$ ), **c** event- and site-corrected ( $\sigma_0$ ) and **d** single-zone ( $\sigma$ ) and total ( $\sigma_{tot}$ ) standard deviation

differences at the short periods, observed in the site effects in Italy and France. The variability of the leftover aleatory residual,  $\sigma_0$ , is closer to that of ITA18, while the values of the total variability reflect the differences seen in terms of between-event standard deviation.  $\sigma_{tot}$  is still quite high but is comparable with the total standard deviation of the Ameri et al. (2017) GMM for low magnitude earthquakes.

## 5.2 Model for VH ratio

### 5.2.1 Functional form

As shown in Fig. 4, no significant bias is observed for  $\delta_0$  (global offset) for the FR20 dataset with respect to the ITA18 model. As a consequence, no global adjustment is necessary to introduce in the VH model for France. However, considering the trend with distance of the event- and site- corrected component  $\delta W_{es}$  at short periods (Fig. 5f), a positive bias has been observed in the near-source region ( $R_{epi} < 15$  km). To model this deviation, we calibrate an adjustment term, that allows to obtain the prediction for France  $Y_{I,FR20}$  as follows:

$$\log_{10} Y_{I.FR20-VH} = \log_{10} Y_{ITA18-VH} + \delta_{VH} \tag{7}$$

where ITA18-VH is the VH model for Italy (Ramadan et al. 2021) and  $\delta_{VH}$  is the corrective factor.  $\delta_{VH}$  is modelled starting from the median value of event- and site- corrected residuals for the observations with  $R_{epi} < 15$  km. Figure 10 shows the median values of  $\delta W_{es}$  as a function of period for  $R_{epi}$  and  $R_{hyp}$ : since the median values are calculated on a limited number of recordings (about 60), the bias does not change smoothly with period but presents many jumps. For this reason, rather than using the averaged value of  $\delta W_{es}$  as model correction, we prefer to linearize the empirical curve, according to the following expression:

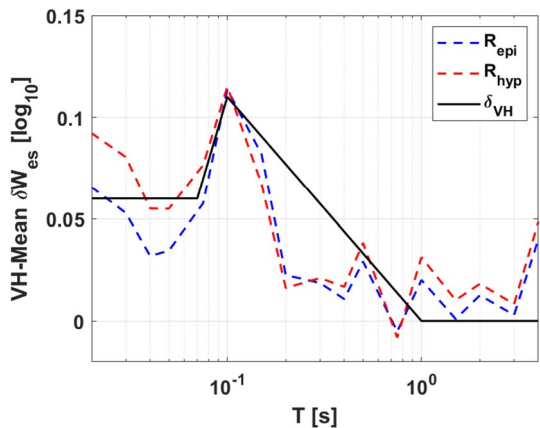
$$\delta_{VH}(T) = \begin{cases} 0.06 & \text{for } T \leq 0.07s \\ 1.667T - 0.056 & \text{for } 0.07s < T \leq 0.1s \\ -0.122T + 0.122 & \text{for } 0.1s < T \leq 1s \\ 0 & \text{for } T > 1s \end{cases} \tag{8}$$

The expression of  $\delta_{VH}$  follows the shape of the mean residual  $\delta W_{es}$  for the first peak at 0.1 s and goes to zero at 1 s; at longer periods we ignore the increasing trend as the long periods of small earthquakes are poorly sampled (see Fig. 10).

### 5.2.2 Model variability

As shown in Fig. 6, the variabilities of the I.FR20-VH are small and similar to the ITA18-VH ones. The introduction of the adjustment term will not change the results, given the small number of data at short distances. Consequently, in this study, the VH standard deviations of the I.FR20 model are assumed to be equal to those of the ITA18 model. VH standard deviations should be used together with the horizontal model variabilities to perform a propagation of uncertainty and obtain the vertical component standard deviations. With respect to the horizontal component, the effect of the VH ratio on the variability is very small and the vertical component standard deviation matches with the horizontal ones (see Ramadan et al. 2021). For this reason and for the sake of simplicity, the vertical component standard deviations are proposed to be equal to the I.FR20 horizontal ones.

**Fig. 10** Mean value of the ITA18 VH event- and site- corrected residuals ( $\delta W_{es}$ ) as a function of period for the records within 15 km and  $\delta_{VH}(T)$  model to correct ITA18 VH in near-source condition for French events





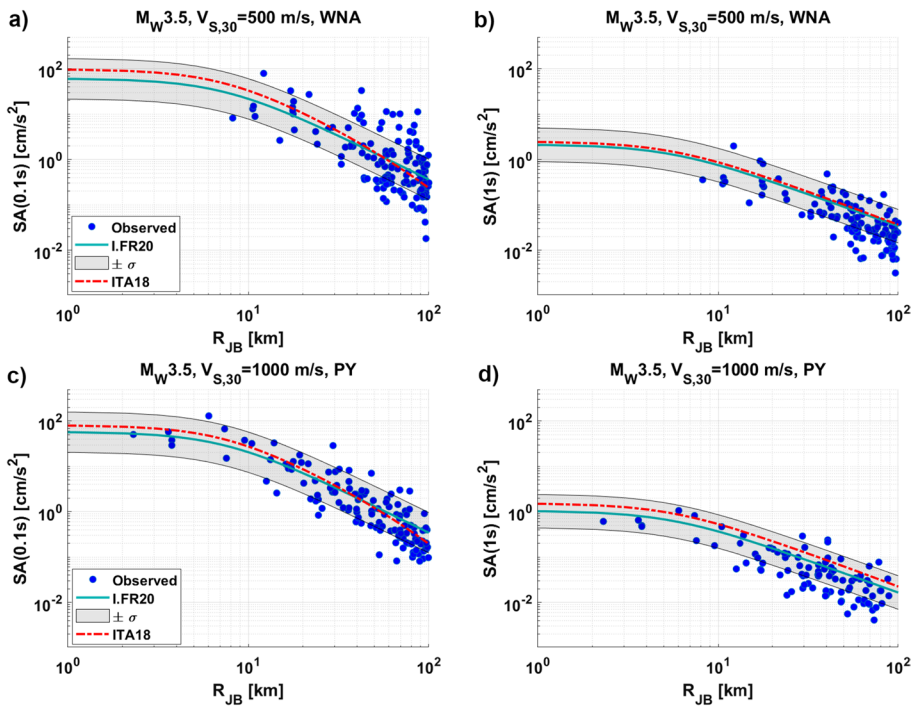
## 6 Predicted spectra for France

### 6.1 Horizontal ground motion

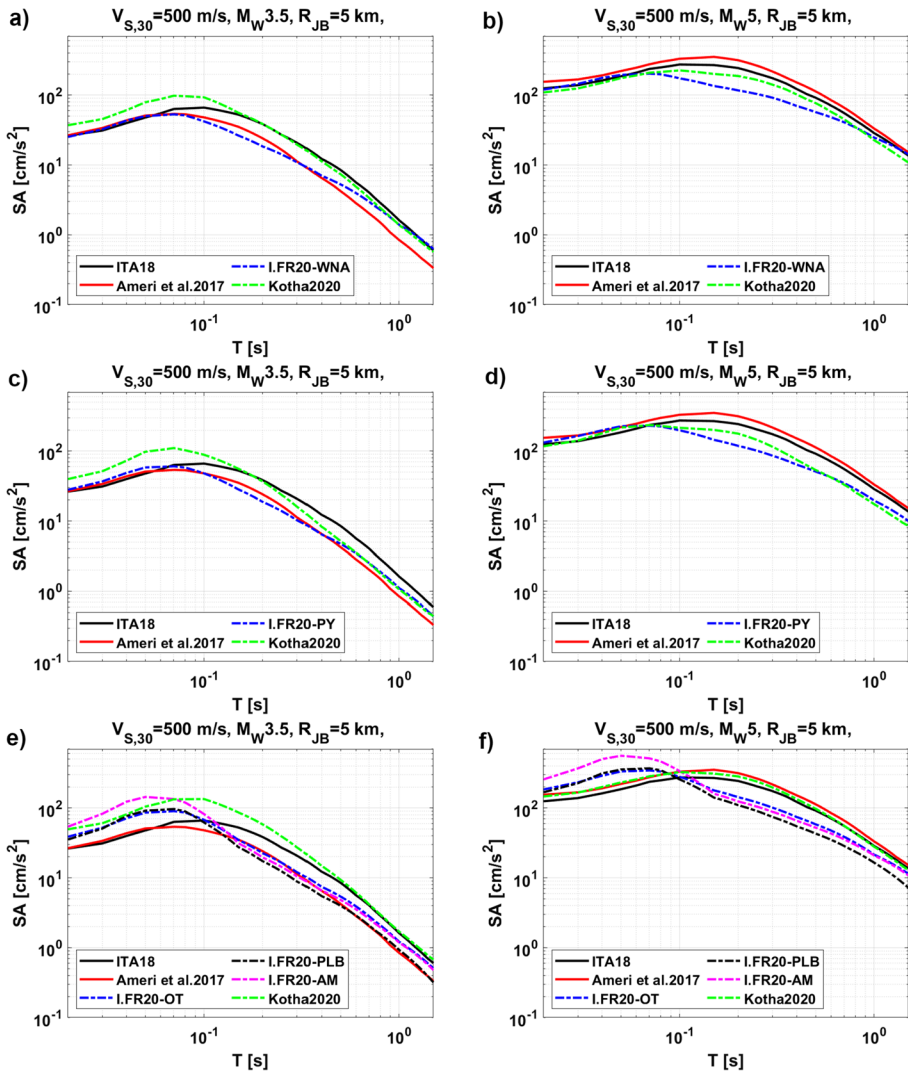
In this section the predictions of the proposed model for horizontal components are presented in comparison with the reference model ITA18. Figure 11 shows the model predictions as a function of distance up to 100 km by the I.FR20 (this study) and ITA18 for two different scenarios ( $M_w$  3.5 and  $V_{S,30}=500$  m/s for WNA;  $M_w$  3.5 and  $V_{S,30}=1000$  m/s for PY), selected in order to display in the figure the available observations in near source conditions.

The above figure demonstrates that the observed data trend is quite well predicted by the proposed model, with most of the observations included in the associated standard deviation. However, in some cases such as the WNA region (Western and Northern Alps), there is a high variability among the records, and some of them at long distances exceed the median prediction plus the standard deviation. We can clearly observe the slower attenuation of the proposed model with respect to the predictions of ITA18.

Figure 12 shows predicted spectra by the proposed model for the 5 different source regions and by the ITA18, Ameri et al. (2017) and Kotha et al. (2020) GMMs at short distance ( $R_{JB}=5$  km) for  $V_{S,30}=500$  m/s and two different magnitudes ( $M_w=3.5$  and



**Fig. 11** GMMs predictions for horizontal component (this study and ITA18) and observed records as a function of Joyner-Boore distance for two spectral ordinates at  $T=0.1$  s and 1 s and a scenario with  $M_w=3.5$  and sites with  $V_{S,30}=500$  m/s located in WNA (a and b) and sites with  $V_{S,30}=1000$  m/s located PY (c and d). The observed data are selected considering  $M_w \pm 0.3$  and  $V_{S,30} \pm 300$  m/s



**Fig. 12** Horizontal acceleration response spectra predicted by I.FR20, ITA18, Ameri et al. (2017) and Kotha et al. (2020) GMMs for two scenarios with  $M_w$  3.5 (a, c and e) and  $M_w$  5.0 (b, d and f),  $R_{JB} = 5$  km and  $V_{S,30} = 500$  m/s. Top panel (a and b): GMMs predictions for WNA; middle panel (c and d): GMMs predictions for PY; bottom panel: GMMs predictions for the other zones (OT, AM and PLB)

$M_w = 4.5$ ). Electronic supplement #1 also provide the plot (see Figure A4) for longer distances ( $R_{JB} = 50$  km).

The structure of the model by Kotha et al. (2020) is typical of the non-ergodic GMMs (Al Atik et al. 2010) and is similar to the corrective model we propose for France (Eq. (5)) since they both provide regional non-ergodic terms to account for differences in ground motion behaviour among different source and attenuation regions. The zonations used by Kotha et al. (2020) and those proposed in this paper are different; therefore, it was necessary to identify in Kotha’s zonation those that could roughly match ours. Figure A5

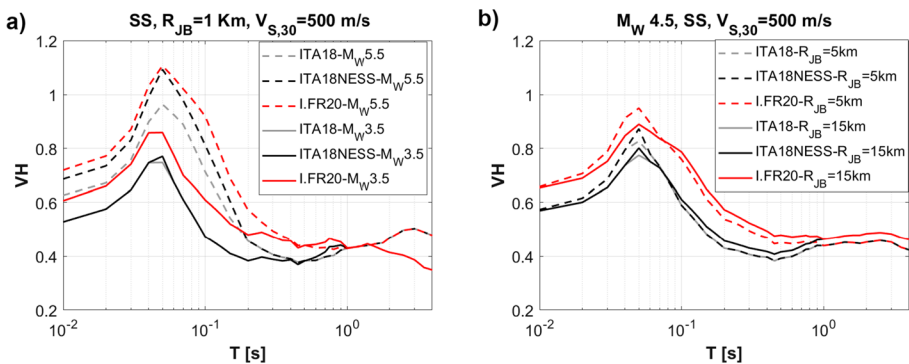
in the electronic supplement #1 shows the source zones and attenuation zones of the Kotha et al. (2020) model, used for the predictions in Fig. 12: we can clearly identify zones corresponding to WNA and PY for both source and attenuation corrections proposed by Kotha et al. (2020). For the other zones (AM, PLB and OT), we consider the Kotha et al. (2020) predictions without regional corrections.

The response spectra predicted by I.FR20 show a peak at shorter periods (between  $T=0.04-0.06$  s, depending on the region) than the ITA18, Ameri et al. (2017) and Kotha et al. (2020), which show the peak at periods about  $T=0.1-0.15$  s. The I.FR20 SA amplitudes for WNA and PY are the lowest, compared to the other GMMs and to other regions, for both  $M_W$  3.5 and 5.0. This result supports the suitability of the source regionalization, since the contribution of  $\delta LL_z$  counterbalances the excess of high frequencies that was observed in the global bias (Fig. 4), resulting in ground motion values equal or lower than the reference GMM for these regions. In contrast, the spectral amplitudes of I.FR20 are higher than those predicted by the other models for PLB, AM and OT regions at short periods (up to  $T=0.1$  s) and become lower for longer periods. In these cases, the increase in high frequency content highlighted by the global bias (Fig. 4) is enhanced. This result suggests that such larger high frequency content might be representative of areas in mainland France with rarer seismicity.

For  $M_W$  3.5 (Fig. 12a, b) the I.FR20 spectra of PY and WNA regions show similar predictions to the Ameri et al., (2017) model; while, for  $M_W$  5.0, I.FR20 predictions are lower than those of the other models for most periods ( $T > 0.07$  s). Regarding the model of Kotha et al. (2020), a different scaling with magnitude is observed with respect to I.FR20 (and thus also ITA18). Indeed, for  $M_W$  3.5, the predictions are higher, while for  $M_W$  5.0 the predicted values are lower and closer to those of the GMM calibrated in this study.

### 6.2 VH ratio

The spectral predictions of VH ratio are shown for two different scenarios in Fig. 13 and the results are compared with the reference model ITA18-VH and ITA18-VH-NESS (Ramadan et al. 2021), the latter being the model for Italy, corrected for near-source effects. The comparison is presented for different  $M_W$  and  $R_{JB}$  scenarios with strike-slip fault and  $V_{S,30} = 500$  m/s. Note that the scenario  $M_W = 5.5$  is beyond the domain of applicability of



**Fig. 13** VH prediction scenarios in near-source for I.FR20-VH (this study), ITA18-VH and ITA18-VH-NESS (Ramadan et al. 2021) for different **a** magnitudes and **b** Joyner-Boore distances

the corrective model I.FR20-VH, and that the comparison is only made for discussion purposes, as the NESS validity domain only begins at  $M_w = 5.5$ .

Figure 13a shows that, for  $M_w \geq 5.5$ , I.FR20 predictions (red line) are in agreement with the ITA18-NESS ones (black solid line), confirming the validity of the approach adopted in this study. For smaller events ( $M_w < 5.5$ ), the I.FR20 model shows a higher VH ratio with respect to ITA18 and ITA18-VH-NESS, because the correction coefficient in ITA18-VH-NESS is significant only for  $M_w$  higher than 5. With respect to  $R_{JB}$  (Fig. 13b), we observe higher I.FR20-VH ratios with respect to the other models, since the correction coefficient for the French dataset is constant over distance.

## 7 Conclusions

France has a peculiar seismicity, characterized by events of small-to-medium magnitude and lower frequency of occurrence than other tectonically active European regions, such as Greece, Turkey, and Italy. In addition, seismicity is spatially distributed over the territory in an inhomogeneous way, with high earthquake concentration in some areas (along the Alpine and Pyrenean chains) where it is more frequent, and with sparse occurrence in others, where it is much rarer (Armorican Massif, central France and the Parsi and Aquitaine basins).

These characteristics make it rather complex to determine an indigenous GMM for seismic hazard purposes, relying only on records of events that have occurred in the area under consideration. In fact, the few attempts to determine an empirical model for France have always benefited from the introduction of data from other regions, such as the GMM proposed by Ameri et al. (2017).

In this paper, to overcome the paucity of data, we develop a GMM for continental France by adapting an existing active shallow crustal model to a conveniently constructed dataset of French ground motion data. The proposed model is valid for PGA, PGV and the ordinates of the response spectrum in the period interval  $T=0.01-2$  s, both for the horizontal component and the vertical-to-horizontal ratio. The procedure consists of the following steps:

1. A dataset of recordings is prepared for the area under investigation. In our case, the dataset, called FR20, is a subset of the database of Traversa et al. (2020) and it consists of 2327 records from 290 earthquakes recorded by 121 stations during the period 1996–2019 in the  $M_w$  range 3.0–5.2;
2. A reference GMM is identified, related to a neighbouring region and/or one with (partially) similar seismic motion characteristics. In this case, the ITA18 model is considered, i.e., the most recent GMM for the Italian territory, available for both the horizontal component (Lanzano et al. 2019a) and the vertical-to-horizontal (VH) ratio (Ramadan et al. 2021);
3. The residuals between the empirical dataset and the reference GMM are calculated to identify the presence of systematic biases on the total residual and with respect to the scaling with the main explanatory variables. In the case of FR20, the horizontal residuals show a total bias over the whole range of periods, leading to an underprediction of ground motion at short periods ( $T=0.01-0.2$  s) and a moderate overprediction at longer periods. In addition, the attenuation with distance is slower than that predicted by ITA18

- at short periods. The VH residuals, on the other hand, show that ITA18 only tends to underpredict the amplitudes of the FR20 data near the source ( $R_{\text{epi}} < 15$  km);
4. Based on the values of the standard deviation associated with the residuals, the benefit of adding parameters/dependencies or random effects to the functional form of the reference GMM is evaluated. In the case of FR20, the variability of the between-event residual ( $\tau$ ) is particularly high for the horizontal component and can be partially reduced by accounting for systematic differences among the observed ground motion amplitudes for events originating in different source areas (e.g., differences in ground motion produced by earthquakes in the Alps and in the Armorican Massif);
  5. Finally, a corrective model is calibrated that accounts for all observed biases and additional terms useful for improving the predictive capability of the regional model and reducing the associated variability. The correction models proposed for France are those in Eq. (5) for the horizontal component ( $\delta_H$ ) and in Eq. (8) for the VH ( $\delta_{VH}$ ).

The predictions corrected for  $\delta_H$  and  $\delta_{VH}$  aim at capturing the main characteristics of ground motion more accurately, showing peculiar spectral shapes for different areas of the France territory. As a matter of fact, the regionalization of the source effect results in significantly different adjustment terms for the different regions, which either compensate or enhance the excess of high frequencies observed over the whole FR20 (Fig. 4). This suggests that such rich high-frequency content in the response spectra might be peculiar of the more tectonically stable areas of metropolitan France (Armorican Massif, Central France), while it vanishes in the more tectonically active areas (Pyrenees, Alps).

Despite the good performance of the median prediction, the model still has a large associated variability, especially for the between-event term. This evidence was already pointed out by other authors (Ameri et al. 2017; Kotha and Traversa 2024) and might possibly be related to a lower quality of the source metadata in France, which might need to be revised in the future.

The validity limits of the correction-terms are closely related to those of the reference dataset, especially in terms of the Mw range, which varies between 3.0 and 5.2. Extension of the model to higher magnitudes for seismic hazard applications needs to be addressed by introducing other constraints, which, in the future, could be provided by the results of physics-based 3D numerical simulations (Smerzini et al. 2023) or other records of earthquakes in neighbouring areas of medium-to-low seismicity (e.g., Belgium, Germany, Switzerland, Spain, etc.), for which ground motion is expected to have similar characteristics. However, the similarity between ITA18 and the ground motion observed in the Alps and Pyrenees chains suggests that the adjusted model could be tested for hazard analysis by extending the range of validity in magnitude to higher values. For the less sampled zones, the high frequency excess could be interpreted by the higher earthquake stress drops expected for these zones, due to the immaturity of the faults (Radiguet et al. 2009).

In future, more complex functional for  $\delta_H$  and  $\delta_{VH}$  could be considered, with additional parametrization and random-effects. Based on the results of this study, some hints to move in this direction are:

- Besides obviously increasing the dataset as new data become available, it is crucial to constantly improve the quality of event, site and recording metadata as much as possible. In all the strong motion datasets, the most problematic event metadata are

magnitude and depth, which can have a considerable impact on the final prediction of the model; the available earthquake catalogues often contain unharmonized magnitude estimates (using different methods, conversions), while depth is often assessed in a conventional manner due to the limitations of the monitoring system. The station metadata could be also improved, for example determining  $V_{S,30}$  using Vs profiles from geophysical measurements, or determining additional site parameters would also allow to model the differences between sites in France and Italy in the ITA18 adjustment, especially with regard to the characteristics of stations on outcropping rock (Cotton et al. 2006);

- The regionalisation of sources can be improved by introducing an *ad-hoc* zonation, based on the spatial pattern of event residuals (see Brunelli et al. 2023), rather than using an existing one built for different purposes. Seismological parameters, such as stress drop, which have been shown to capture the variability of motion at high frequencies, could also be introduced in an advanced version of the model (Ameri et al. 2017; Sgobba et al. 2023);
- Finally, to complete the regionalisation of ground motion in France, it would be appropriate to study the local characteristics of the source-to-site propagation. As a matter of fact, Sung et al. (2023) observed that there is a strong variability of the attenuation across the French territory. Introducing a spatially variable  $\delta c_3$  term that corrects for the GMM attenuation with distance is not straightforward and depends heavily on the number of paths sampled from the dataset. However, it represents one of the fundamental aspects to be explored in the future to interpret the variability of ground motion in France.

## Appendix 1

### Summary of ITA18 GMMs

The functional form for horizontal component, ITA18-H, is:

$$\log_{10}Y = a + F_M(M_W, SoF) + F_D(M_W, R) + F_S(V_{S30}) + \sigma \tag{A1}$$

$$F_M(M_W, SoF) = f_j SoF_j + \begin{cases} b1(Mw - Mh) \text{ where } Mw \leq Mh \\ b2(Mw - Mh) \text{ where } Mw > Mh \end{cases} \tag{A2}$$

$$F_D(M_W, R) = [c_1(M - M_{ref}) + c_2] \log_{10} \sqrt{R^2 + h^2} + c_2 \sqrt{R^2 + h^2} \tag{A3}$$

$$F_S(V_{S30}) = k \left( \frac{V_0}{800} \right) \begin{cases} V_0 = V_{S30} \text{ for } V_{S30} \leq 1500 \text{ m/s} \\ V_0 = 1500 \text{ m/s for } V_{S30} > 1500 \text{ m/s} \end{cases} \tag{A4}$$

$$\sigma = \sqrt{\tau + \phi_{S2S} + \phi_0} \tag{A5}$$

Y is the observed IM, i.e. 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. 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 non-redundant 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.

The hinge magnitude  $M_h$ , the reference magnitude  $M_{ref}$  and the pseudo-depth  $h$  are obtained in a first step non-linear regression. The coefficients  $a$ ,  $b_1$ ,  $b_2$ ,  $c_1$ ,  $c_2$ ,  $c_3$ ,  $k$  and  $f_j$  ( $f_1$  for strike-slip,  $f_2$  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, in order to estimate the partially non-ergodic sigma according to Al Atik et al. (2010), where  $\tau$  and  $\phi_{S2S}$  represent between-event and site-to-site variability, respectively, and  $\phi_0$  is the standard deviation of the event- and site- corrected residuals. The coefficients were provided in Lanzano et al. (2019a).

The general mathematical expression for vertical to horizontal ratio (ITA18-VH) is the same of Eq. [A1] and the parameters  $Y$  for which the model was calibrated are the same as for the horizontal component.

The functions related to source and attenuation are different:

$$F_M(M_w, SoF) = bM_w + f_j SoF_j \quad (A6)$$

$$F_D(M_w, R) = [c_1(M - M_{ref}) + c_2] \log_{10} \sqrt{R^2 + h^2} \quad (A7)$$

Scaling with  $V_{S30}$  (Eq. [A4]) and residual decomposition (Eq. [A5]) are unchanged from the model for the horizontal component. The coefficients were provided in Ramadan et al. (2021).

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10518-024-01890-4>.

**Acknowledgements** The authors would like to thank the manuscript reviewers, Gabriele Ameri and Kris Vanneste, for their helpful comments.

**Author contributions** All the authors equally contributed to the analysis of the data and the writing of the manuscript.

**Funding** Open access funding provided by Istituto Nazionale di Geofisica e Vulcanologia within the CRUI-CARE Agreement. This work is the result of an agreement between the EDF (Electricité DE France) company and the INGV (Istituto Nazionale di Geofisica e Vulcanologia) of Milan, with the purpose of characterising the seismic ground motion in near-source condition for low-to-moderate seismic areas.

**Data availability** The database, from which the FR20 dataset is extracted, is accessible as an electronic supplement to the publication by Traversa et al. (2020).

## Declarations

**Competing interests** The authors declare that they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article

are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Akkar S, Sandikkaya MA, Şenyurt M, Azari Sisi A, Ay BÖ, Traversa P, Godey S (2014) Reference database for seismic ground-motion in Europe (RESORCE). *Bull Earthq Eng* 12:311–339
- Al AL, Abrahamson N, Bommer JJ, Scherbaum F, Cotton F, Kuehn N (2010) The variability of ground-motion prediction models and its components. *Seismol Res Lett* 81(5):794–801
- Ameri G, Drouet S, Traversa P, Bindi D, Cotton F (2017) Toward an empirical ground motion prediction equation for France: accounting for regional differences in the source stress parameter. *Bull Earthq Eng* 15:4681–4717
- Atkinson GM (2008) Ground-motion prediction equations for eastern North America from a referenced empirical approach: Implications for epistemic uncertainty. *Bull Seismol Soc Am* 98(3):1304–1318
- Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw* 67(1):1–48
- Bayless J, Abrahamson NA (2019) Summary of the BA18 ground-motion model for Fourier amplitude spectra for crustal earthquakes in California. *Bull Seismol Soc Am* 109(5):2088–2105
- Bindi D, Pacor F, Luzi L, Puglia R, Massa M, Ameri G, Paolucci R (2011) Ground motion prediction equations derived from the Italian strong motion database. *Bull Earthq Eng* 9:1899–1920
- Brunelli G, Lanzano G, Luzi L, Sgobba S (2023) Data-driven zonation for modelling the regional source and propagation effects into Ground Motion Models in Italy. *Soil Dyn Earthq Eng* 166:107775
- Cara M, Cansi Y, Schlupp A, Arroucau P, Béthoux N, Beucler E, Van Der Woerd J (2015) SI-Hex: a new catalogue of instrumental seismicity for metropolitan France. *Bulletin De La Société Géologique De France* 186(1):3–19
- Causse M, Cornou C, Maufroy E, Grasso JR, Baillet L, El Haber E (2021) Exceptional ground motion during the shallow Mw 4.9 2019 Le Teil earthquake. *France. Communications Earth & Environment* 2(1):1–9
- CEN - European Committee for Standardization (2004). Eurocode 8. Design of structures for earthquake resistance - Part 1: General rules, seismic actions and rules for buildings, EN 1998–1.
- Cornou C, Ampuero J-P, Aubert C, Audin L, Weng H (2021) Rapid response to the Mw 4.9 earthquake of November 11, 2019 in Le Teil, Lower Rhone Valley, France. *Comptes Rendus Geoscience Sciences De La Planete* 353:441–463
- Cotton F, Scherbaum F, Bommer JJ, Bungum H (2006) Criteria for selecting and adjusting ground-motion models for specific target regions: Application to central Europe and rock sites. *J Seismolog* 10:137–156
- Drouet S, Cotton F (2015) Regional stochastic GMPEs in low-seismicity areas: Scaling and aleatory variability analysis—Application to the French Alps. *Bull Seismol Soc Am* 105(4):1883–1902
- Drouet S, Cotton F, Guéguen P (2010)  $V_{S,30}$ ,  $\kappa$ , regional attenuation and Mw from accelerograms: application to magnitude 3–5 French earthquakes. *Geophys J Int* 182(2):880–898
- Drouet S, Ameri G, Le Dortz K, Secanell R, Senfaute G (2020) A probabilistic seismic hazard map for the metropolitan France. *Bull Earthq Eng* 18:1865–1898
- Duverger C, Mazet-Roux G, Bollinger L, Trilla AG, Vallage A, Hernandez B, Cansi Y (2021) A decade of seismicity in metropolitan France (2010–2019): the CEA/LDG methodologies and observations. *Bulletin de la Société Géologique de France*, 192(1).
- Forte G, Chioccarelli E, De Falco M, Cito P, Santo A, Iervolino L (2019) Seismic soil classification of Italy based on surface geology and shear-wave measurements. *Soil Dyn Earthq Eng* 122:79–93
- Kotha SR, Weatherill G, Bindi D, Cotton F (2020) A regionally-adaptable ground-motion model for shallow crustal earthquakes in Europe. *Bull Earthq Eng* 18(9):4091–4125
- Kotha SR, Traversa P (2024) A Bayesian update of Kotha et al. (2020) ground-motion model using Résif dataset. *Bull Earthq Eng* 1–27.
- Lanzano G, Luzi L, Pacor F, Felicetta C, Puglia R, Sgobba S, D'Amico M (2019a) A revised ground-motion prediction model for shallow crustal earthquakes in Italy. *Bull Seismol Soc Am* 109(2):525–540
- Lanzano G, Sgobba S, Luzi L, Puglia R, Pacor F, Felicetta C, Bindi D (2019b) The pan-European Engineering Strong Motion (ESM) flatfile: compilation criteria and data statistics. *Bull Earthq Eng* 17:561–582



- Lanzano G, Ramadan F, Luzi L, Sgobba S, Felicetta C, Pacor F, D'Amico M, Puglia R, Russo E (2022) Parametric table of the ITA18 GMM for PGA, PGV and Spectral Acceleration ordinates. Istituto Nazionale di Geofisica e Vulcanologia (INGV). Available at: [https://doi.org/10.13127/ita18/sa\\_flatfile](https://doi.org/10.13127/ita18/sa_flatfile)
- Luzi L, Pacor F, Puglia R, Lanzano G, Felicetta C, D'Amico M, Chioccarelli E (2017) The central Italy seismic sequence between August and December 2016: Analysis of strong-motion observations. *Seismol Res Lett* 88(5):1219–1231
- Mayor J, Traversa P, Calvet M, Margerin L (2018) Tomography of crustal seismic attenuation in Metropolitan France: implications for seismicity analysis. *Bull Earthq Eng* 16(6):2195–2210
- Mazzotti S, Aubagnac C, Bollinger L, Oscanoa KC, Delouis B, Do Paco D, Trilla A (2021) FMHex20: An earthquake focal mechanism database for seismotectonic analyses in metropolitan France and bordering regions. *Bulletin De La Société Géologique De France*, 192(1).
- Mori F, Mendicelli A, Moscatelli M, Romagnoli G, Peronace E, Naso G (2020) A new vs30 map for Italy based on the seismic microzonation dataset. *Engineering Geology*, 275.
- Pacor F, Felicetta C, Lanzano G, Sgobba S, Puglia R, D'Amico M, Iervolino I (2018) NESS1: A worldwide collection of strong-motion data to investigate near-source effects. *Seismol Res Lett* 89(6):2299–2313
- Radiguet M, Cotton F, Manighetti I, Campillo M, Douglas J (2009) Dependency of nearfield ground motions on the structural maturity of the ruptured faults. *Bull Seismol Soc Am* 99(4):2572–2581
- Ramadan F, Smerzini C, Lanzano G, Pacor F (2021) An empirical model for the vertical-to-horizontal spectral ratios for Italy. *Earthquake Eng Struct Dynam* 50(15):3937–4219
- RESIF (1995) RESIF-RAP French Accelerometric Network . RESIF - Réseau Sismologique et géodésique Français. Available at: <https://doi.org/10.15778/RESIF.RA>
- Ritz JF, Baize S, Ferry M et al. (2020) Surface rupture and shallow fault reactivation during the 2019 Mw 4.9 Le Teil earthquake, France. *Commun Earth Environ* 1, 10. <https://doi.org/10.1038/s43247-020-0012-z>.
- Rodriguez-Marek A, Rathje EM, Bommer JJ, Scherbaum F, Stafford PJ (2014) Application of single-station sigma and site-response characterization in a probabilistic seismic-hazard analysis for a new nuclear site. *Bull Seismol Soc Am* 104(4):1601–1619
- Sabetta F, Pugliese A (1987) Attenuation of peak horizontal acceleration and velocity from Italian strong-motion records. *Bull Seismol Soc Am* 77(5):1491–1513
- Sgobba S, Lanzano G, Pacor F (2021) Empirical nonergodic shaking scenarios based on spatial correlation models: An application to central Italy. *Earthq Eng Struct Dynam* 50(1):60–80
- Sgobba S, Lanzano G, Colavitti L, Morasca P, D'Amico MC, Spallarossa D (2023) Physics-based parametrization of a FAS nonergodic ground motion model for Central Italy. *Bull Earthq Eng* 1–27.
- Smerzini C, Vanini M, Paolucci R et al (2023) Regional physics-based simulation of ground motion within the Rhône Valley, France, during the MW 4.9 2019 Le Teil earthquake. *Bull Earthq Eng* 21:1747–1774. <https://doi.org/10.1007/s10518-022-01591-w>
- Stafford PJ (2014) Crossed and nested mixed-effects approaches for enhanced model development and removal of the ergodic assumption in empirical ground-motion models. *Bull Seismol Soc Am* 104(2):702–719
- Sung CH, Abrahamson NA, Kuehn NM, Traversa P, Zentner I (2023) A non-ergodic ground-motion model of Fourier amplitude spectra for France. *Bull Earthq Eng* 21(11):5293–5317
- Traversa P, Maufroy E, Hollender F, Perron V, Bremaud V, Shible H, Douste-Bacque I (2020) RESIF RAP and RLBP dataset of earthquake ground motion in mainland France. *Seismol Res Lett* 91(4):2409–2424