

A Bayesian Parametric Ground-Motion Prediction Model: Response Spectra of Shallow Crustal Seismicity in Europe

Pauline Georges

`pauline.georges@univ-grenoble-alpes.fr`

Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, Univ. Gustave Eiffel, ISTerre

Sreeram Reddy Kotha

Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, Univ. Gustave Eiffel, ISTerre

Emmanuel Chaljub

Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, Univ. Gustave Eiffel, ISTerre

Research Article

Keywords: Ground-motion model, Response spectra, Bayesian regression, Attenuation regionalisation, Seismic hazard and risk, Europe

Posted Date: October 7th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10852247/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

A Bayesian Parametric Ground-Motion Prediction Model: Response Spectra of Shallow Crustal Seismicity in Europe

Pauline Georges^{1*†}, Sreeram Reddy Kotha^{1†} and
Emmanuel Chaljub^{1†}

¹Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, Univ.
Gustave Eiffel, ISTerre, Grenoble, France.

*Corresponding author(s). E-mail(s):

pauline.georges@univ-grenoble-alpes.fr;

Contributing authors: sreeram-reddy.kotha@univ-grenoble-alpes.fr;

emmanuel.chaljub@univ-grenoble-alpes.fr;

[†]These authors contributed equally to this work.

Abstract

Ground-Motion Prediction Models (GMPM) predict log-normal distributions of ground-motion at a location as a linear combination of source, path, and site parametric functions. The linear-viscoelastic assumption allows calibrating path functions using abundant small-moderate sized earthquake data and using them for rare large earthquakes. As wave propagation is a complex phenomenon involving multiple mechanisms that vary considerably across low and high seismicity regions of Europe, it is necessary that path effects are regionally calibrated. Therefore, we combine three publicly available datasets of 5% damped horizontal response spectra of shallow crustal earthquakes in France, Belgium, and high seismicity regions of Europe, and regress a GMPM. The model supersedes the previous pan-European regionally adaptable GMPM in three aspects: 1) Broader range of homogenised moment-magnitudes ($2 < M < 8$); 2) Event and site location dependent attenuation that better reflects the physics of wave propagation; 3) Bayesian regression that allows for subsequent updates. The epistemic and aleatory uncertainties are notably smaller, and information criteria confirm a superior predictive performance compared to the previous pan-European GMPM.

Keywords: Ground-motion model, Response spectra, Bayesian regression, Attenuation regionalisation, Seismic hazard and risk, Europe

1 Introduction

Ground-Motion Prediction Models (GMPMs) are typically parametric models that summarise and predict the (dependent) ground-motion observations as a function of (independent) predictor parameters. The predicted ground-motions at a location are often Gaussian distributions $\mathcal{N}(\mu, \sigma^2)$; with μ as the mean/median and σ as the standard-deviation. The $\mu = G(\mathbf{M}, R, \theta)$ are parametric functions of event moment-magnitude $\mathbf{M}$ (Hanks and Kanamori 1979), event-to-site distance R , and site-response parameter θ . The natural randomness of observed ground-motions not modelled into $G(\mathbf{M}, R, \theta)$ is quantified in to the *apparent* aleatory variability σ . Such parametrisation structures the GMPMs and constrains their behaviour to be physically meaningful when interpolating and extrapolating beyond the scenarios $[\mathbf{M}, R, \theta]$ sampled in their calibration ground-motion datasets or *flatfiles*. This is a clear advantage of parametric GMPMs over the non-parametric models (e.g. Random Forests), which despite their flexibility and apparently smaller misfit to data (Trugman and Shearer 2018), are not applicable beyond their training dataset.

Aside from the unexplained σ , recent studies have practiced reporting also the within-model epistemic uncertainty σ_μ (Atik and Youngs 2014; Kotha et al. 2020); which is an $[\mathbf{M}, R, \theta]$ dependent uncertainty on the μ . σ and σ_μ depend on the complexity of $G(\mathbf{M}, R, \theta)$ and the uncertainty/errors on $[\mathbf{M}, R, \theta]$ measurements in the flatfile (Laurendeau and Kotha 2025; Kuehn et al. 2020). Liou and Abrahamson (2024) suggested reporting also the epistemic uncertainty on σ given the increasingly complex parametrisation of GMPMs. This is especially true with regards to the shift from ergodic to non-ergodic GMPMs, which requires evaluating the statistical significance of the changes in σ and σ_μ from increased complexity of $G(\mathbf{M}, R, \theta)$ calibrated on large datasets of *poor* quality $[\mathbf{M}, R, \theta]$ metadata.

Over the last decade, the shift from ergodic to non-ergodic GMPMs focused primarily on resolving the ergodicities via the so-called site-specific GMPMs; with site-specific adjustments quantified and applied to μ to make site-specific predictions. Fundamental to non-ergodic GMPMs is the *repeatability* of specific physical effects. In this sense, site-responses have shown to remain stationary in time (with some exceptions), allowing them to be quantified from past data and applied to future earthquakes, while removing their site-to-site response variability from σ . As a result, site-specific GMPMs featured a significantly smaller σ than their ergodic versions; but with a significant increase in σ_μ for the poorly sampled sites. It is to be noted that site-specific GMPMs have become possible thanks to the rapid expansion and operation of high-quality seismic networks, and the open-access to automatically processed flatfiles (Zaccarelli et al. 2019; Mascandola and Sgobba 2025). The newer networks could assimilate high-quality broadband data from numerous small-moderate sized earthquakes of $\mathbf{M} \leq 4.5$, which are essential in constraining site-specific adjustments of hundreds of sites in the seismically active regions.

A critical aspect of parametric non-ergodic GMPMs is the assumptions underpinning their functional forms and the data distributions, which are adopted from the classical elastodynamic theory. Assuming linear-viscoelastic behaviour of the propagation medium under infinitesimal strain, the observed surface ground-motion in time-domain can be expressed as a convolution of separated variables: event (E) Source Time Function, and path (P) and site (S) Green's functions. In the frequency domain, the ground-motion Fourier spectrum is a linear product of event source spectrum, and path and site transfer functions. Logarithm of the observed peak ground-motion $\ln(Y_{e,s})$ at site s located at distance $R_{e,s}$ from an event e of $\mathbf{M}_e$, as is the case in GMPMs, would then become a linear addition of event, path, and site parametric functions: $\ln(Y_{e,s}) = G_E(\mathbf{M}_e) + G_P(\mathbf{M}, R_{e,s}) + G_S(\theta) + \epsilon_{e,s}$; where $\epsilon_{e,s}$ is the event-and-site specific ground-motion residual from the Gaussian distribution $\mathcal{N}(0, \sigma^2)$. This assumption allows calibration of $G_P(\mathbf{M}, R)$ and $G_S(\theta)$ from abundant small earthquakes - causing infinitesimal strains in a linear-viscoelastic medium - in a target region, and use them in-tandem with $G_E(\mathbf{M})$ calibrated on rare, large earthquake data from an entirely different host region; and often for prospective $\mathbf{M}$ larger than those observed in either regions. Although small earthquake data is indispensable in the shift from ergodic to non-ergodic GMPMs, it is arguable whether their $\mathbf{M}$ and location are accurate, and if the non-ergodic GMPMs calibrated on small earthquake data can be used in predictions for large earthquakes.

It is to be noted that, GMPMs developed under the linear-assumption are not guaranteed to be applicable for hazard critical scenarios in both low and high seismicity regions, i.e., near-source distances of large earthquakes with strains large enough to render both path and site behavior to be nonlinear, and when the linear combination of separated variables becomes invalid. Nevertheless, in low-moderate seismicity regions like France (Walpersdorf et al. 2015) and Belgium, relatively small earthquakes can be damaging to populated valleys and basins, e.g. the $\mathbf{M}4.5$ Le Teil, 2019 in France (Cornou et al. 2021). In addition, large destructive earthquakes are also probable according to the historical archives of the last 800 years, e.g. around 10 events of $\mathbf{M} \geq 5.8$ (Manchuel et al. 2018) in France, and 2 events of $\mathbf{M} \geq 5.8$ in Belgium (Vanneste and Onvani 2025). Probabilistic Seismic Hazard Assessments (PSHA) of low-moderate seismicity regions therefore require GMPMs with $G_P(\mathbf{M}, R)$ and $G_S(\theta)$ calibrated on data from local/regional earthquakes (e.g. $\mathbf{M} \leq 5.5$), and $G_E(\mathbf{M})$ calibrated on large earthquake (e.g. $\mathbf{M} \geq 5.5$) data from high seismicity regions (e.g. Italy, Turkey, Greece, etc.), and a low $\sigma_\mu(4 < M \leq 5)$. This is the so-called *partially* non-ergodic approach to *regionally adaptable* site-specific GMPMs.

The level of non-ergodicity of path effects depends on the complexity of $G_P(\mathbf{M}, R)$. For example, treating each event-to-site path uniquely (i.e. $p = e,s$) requires multiple samples of that specific path. A few studies have produced non-ergodic GMPMs where path effects of multiple paths traversing a 2D grid-cell are averaged as cell-specific path effects (Dawood and Rodriguez-Marek 2013); while a few others averaged the path effects over all paths within broader national administrative boundaries (Gregor

et al. 2014; Sandıkkaya et al. 2025; Kotha et al. 2016; Kale et al. 2015) or geologically and tectonically similar regions with sufficient sites located within them (Kotha et al. 2020). The complexity needs to balance the trade-off between σ and σ_μ , and should reflect known propagation physics. Wave propagation theory and numerical simulations help in appropriately parametrising and calibrating $G_P(\mathbf{M}, R)$ to emulate attenuation physics specific to a region or path; where sufficient amount of data is available to allow meaningful reductions in σ and marginal increases in σ_μ . Once again, the sufficiency of data can be addressed only with inclusion of small earthquakes in the regression flatfiles. Although these small earthquakes themselves may have an insignificant role in hazard assessments, they are indispensable in calibration of partially non-ergodic GMPMs. In this study, the focus is on building upon and refining the regionally adaptable site-specific GMPM of Kotha et al. (2020) ($K20$ from hereon) to produce a new GMPM that can supersede it. There are three major issues identified with $K20$ and addressed by the new GMPM:

- **Applicability in low-moderate seismicity regions:** $K20$ was regressed on the Engineering Strong-Motion dataset (ESM18 by Lanzano et al. (2019) and Bindi et al. (2019)) with data $3 < \mathbf{M} < 7.4$ primarily from Italy, Greece and Turkey. Most of the $\mathbf{M} \leq 5.5$ estimates therein were inferred from empirical $\mathbf{M} - M_L$ relationships, and the $\mathbf{M} > 5.5$ from event-specific studies, as provided in Euro-Mediterranean Earthquake Catalogue (EMEC, Grünthal and Wahlström (2012), Lammers et al. (2019)). As a result, $K20$'s $G_e(\mathbf{M})$ had a rather nonphysical behavior (Kotha et al. 2022b; Fukushima 1996) with a large σ_μ for $\mathbf{M} \leq 4.5$ (Kotha et al. 2022b), limiting its applicability strictly to $\mathbf{M} > 4.5$. In addition, Laurendeau and Kotha (2025) demonstrated that the σ of Kotha et al. (2022a) (and $K20$) are likely inflated by the inhomogeneity and errors in $\mathbf{M} \leq 5.5$ in the ESM-EMEC flatfile. The new GMPM is intend to be well-behaved $G_E(\mathbf{M})$ over a broader $2 < \mathbf{M} < 8$ range, lower σ_μ at $\mathbf{M} \leq 4.5$, and smaller σ than $K20$
- **Regionalisation of apparent anelastic attenuation:** A unique feature of $K20$ with respect to its contemporary GMPMs was that its $G_P(\mathbf{M}, R)$ is regionally adaptable. The *apparent* anelastic attenuation represents the exponential decay of ground-motion with distance due to absorption and scattering of seismic waves along its path in the medium. As it is difficult to separate the two phenomena, we prefix attenuation with the term *apparent*. The apparent anelastic attenuation coefficient within $G_P(\mathbf{M}, R)$ varied with attenuation regions r . The regionalisation model was a Weatherill et al. (2020) modification of the geological-tectonic regionalisation model of Basili et al. (2021) for use in the European Seismic Hazard Maps 2020 (ESHM20, Danciu et al. (2021)). The $K20$ GMPM was able to distinguish regions with faster, average, and slower attenuation within the spatial extent of its regressed dataset. However, as the ESHM20 extended well beyond the spatial coverage of the calibration dataset, most of the low-moderate seismicity regions (e.g. France, Belgium) had no region-specific attenuation adjustments. Subsequently, with the EPOS-FR dataset for France (Traversa et al. 2020), Kotha and Traversa (2024) attempted a Bayesian update of the Kotha et al. (2020) to quantify region-specific adjustments for France. However, although Kotha and Traversa (2024) reported that the high-frequency ground-motion attenuation is slower (or lower) in France compared

to Italy, the *K20* regionalisation was unable to quantify the observed variability within France (Campillo and Plantet 1991; Mayor et al. 2018). In addition, the *K20* regionalisation is independent of the ground-motion frequency and was tied to site location, while attenuation is understood to have a strong frequency dependence and should depend also on event location (Mayor et al. 2018). Therefore, the attenuation regionalisation scheme of Kotha et al. (2020) needed an evaluation and refinement.

- **Inhomogeneity of M :** Small $M \leq 4.5$ earthquake data is indispensable for non-ergodic GMPMs and PSHA in low-moderate seismicity regions. However, their M are often derived from empirical M - M_L conversion formulae, which vary across the regions and are generally calibrated on data from high seismicity regions. *K20* and Kotha et al. (2022a) preferred M from EMEC catalog in their regressions to maintain consistency with M used in seismic source models of ESHM20. But this in turn forced an inconsistency between the events' M and their associated ground-motion observations in ESM18 flatfile. As a result, the GMPMs had σ and σ_μ inflated by the (possible) errors in the predictor parameter M (Laurendeau and Kotha 2025). In this study, particular attention is given to combining three independent flatfiles for low and high seismicity regions, record selection for regression, and preferred M .

In addition to the above improvements, an important change is in the regression algorithm, and estimation of components of σ . While *K20* was regressed using a robust linear mixed-effects algorithm (RLMM, Koller (2016)), the new GMPM is regressed using a Bayesian linear mixed-effects algorithm (INLA, Lindgren and Rue (2015)). This allows the model to be updated with new data, and provides insight on whether the resulting changes in σ and σ_μ are statistically significant.

2 Ground-motion datasets

To develop a regionally adaptable GMPM for shallow crustal earthquakes across low and high seismicity regions of Europe, a compendium of three published, open-access flatfiles is produced in this study. The flatfiles provide 3-component 5% damped spectral accelerations (*SA*) for 36 periods between $T = 0.01 - 10s$, *PGA*, and *PGV*. The GMPM is calibrated to predict the RotD50 combination of the two horizontal component response spectra as defined in Boore (2010).

- **Engineering Strong-Motion dataset (EPOS-ESM):** The ESM18 flatfile used in deriving *K20*, contained data until end of 2016. The updated EPOS-ESM flatfile used in this study contains automatically processed response spectra of $3.2 \leq M \leq 7.7$ observations from 1976 to 2026 occurring in the geographical area 34° to 60° N and -10° to 50° E. 79% of the data are acquired from broadband accelerometers. A second-order acausal time-domain Butterworth filter is applied to the zero-padded acceleration time series and zero-pad removal to make acceleration and displacement consistent after double integration. The usable band-pass frequency range is generally 0.1 - 40.0 Hz. Details on the waveform processing and calculation of response spectra are provided in Lanzano et al. (2019) and Mascandola et al. (2023). This dataset is focused on the high seismicity Euro-Mediterranean regions; with 46% of the data from Italy, 31% from Greece, 12% from Turkey, and the

rest mainly from Switzerland and Romania. Event M are assigned depending on the availability; with first preference to earthquake specific studies, then to the Moment Tensor solutions, and finally, to regional or international bulletins. Revised M from EMEC are also provided in the flatfile. $M \leq 4.5$ constitute 60% of the data in flatfile, near-source data ($R_{epi} \leq 10km$) constitute only 2% of the dataset. Regarding the site-response parameters, about 21% of the shear-wave velocity, e.g. V_{s30} , are from in-situ measurements, while most other sites are provided with Eurocode8 class inferred from geology.

- **EPOS-FR:** The French flatfile is the [Buscetti et al. \(2024\)](#) update of the [Traversa et al. \(2020\)](#) used in [Kotha and Traversa \(2024\)](#), with greater spatial and temporal coverage of the hexagonal France. It provides both response and Fourier spectra of automatically processed waveforms recorded by velocimetric and accelerometric stations in France between 1996 and end of 2021 in the geographical area 41° to 53° N and -6° to 11° E. 60% of the total dataset comes from broadband seismometers. Unlike the EPOS-ESM dataset, the usable frequency range is limited to 0.5–25 Hz to prevent contamination of SAs with poor signal-to-noise ratio at lower frequencies. A waterlevel of 100 dB and a Butterworth passband filter between 0.001 Hz and $0.9f_{Nyquist}$ is applied to correct the instrumental response. The signals are band-pass filtered in the Fourier domain using a fourth-order acausal Butterworth filter. The suggested filter corner frequencies correspond to frequency limits for Signal-to-Noise Ratio ($SNR > 3$). Magnitude are ranging in $1.9 \leq M \leq 5.2$.
- **BELSHAKE:** The Belgian flatfile provides both response and Fourier spectra of earthquakes occurring in 1983-2022 in Belgium and neighbouring countries, recorded by the seismic network operated by Royal Observatory of Belgium. 59% of the data are acquired from short-period seismometers, and 23 % from broadband seismometers. A semi-automated method coded in Python is used to process all waveform data with a subsequent manual correction. The processing steps including detrending, tapering, SNR, minimal and maximal frequencies computation, windowing are elaborated in the [Vanneste and Onvani \(2025\)](#). The earthquake and station meta-data were computed from waveforms available on the Royal Observatory of Belgium waveform archive. The resulting magnitude range is similar to that of EPOS-FR : $1.6 \leq M \leq 5.3$.

2.1 Data selection

The three flatfiles combined contain approximately three-times the number of records used to derive the $K20$ GMPM. 68 % of the data are from EPOS-ESM, 25 % from EPOS-FR, and the rest from BELSHAKE. The number of records, earthquakes and stations used in this regression compared to the previous model is shown in Table 1. The resulting compendium dataset provides a broad coverage of Europe, as shown in Figure 1. The Balkan region is still under-sampled in the compendium flatfile, but the Bayesian GMPM allows subsequent updates when local and regional datasets are made available. However, a more immediate issue with combining independently compiled datasets of overlapping spatial and temporal coverage is to identify duplicates of records and differences in their processing, identifying differences in metadata of shared events and stations, and selection for GMPM regression.

Table 1 Comparison of number of records, sites, events, site attenuation regions, event attenuation regions, tectonic localities in this study against *K20*; and in quantifying the estimates of ϕ_0 , ϕ_{S2S} , τ_0 , $\tau_{c3,s}$, $\tau_{c3,e}$, τ_{L2L} , respectively, shown in Figure 3 at $T=0.1$ and $1s$

T(s)	GMPM	Records ϕ_0	Sites ϕ_{S2S}	Events τ_0	Attenuation (site) $\tau_{c3,s}$	Attenuation (event) $\tau_{c3,e}$	Localities τ_{L2L}
0.10	This study	55 207	3 423	2 624	713	655	574
0.10	K20	18 222	1 829	927	46	0	56
1.00	This study	48 084	3 400	2 365	714	653	534
1.00	K20	18 217	1 829	926	46	0	56

- **Events:** Only shallow tectonic earthquakes with hypocentral depth less than $40km$ or the Moho depth (Grad et al. 2009) at event location, and recorded by at least 3 stations are retained in the regression. EPOS-FR and EPOS-ESM datasets share events at the France-Italy and France-Switzerland national borders in the Western Alps. These events were identified based on their location, origin-time, M and M_L . BELSHAKE dataset identifies the events common with EPOS-FR by origin time difference, M difference, and M_L difference thresholds. Further comparisons of spectral acceleration ratios and location criteria were tested but found to be unnecessary. Only the EPOS-FR event metadata are retained in the compendium dataset.
- **Sites:** All sites are used regardless of their depth, installation type, and availability of site-response parameters, because the site-responses are quantified into site-specific random-effects explained subsequently.
- **Records:** In case of records shared by the BELSHAKE, EPOS-FR and EPOS-ESM datasets, only those from EPOS-FR are retained in the compendium dataset because EPOS-FR apply a higher-order filter (4^{th} order), which is more suitable for the small-moderate earthquake records likely to be contaminated by noise at low frequencies. GMPM regressions are performed for each T without considering cross-correlation of $SA(T)$. Therefore, for each T , only those records with $T \leq 0.8/f_{hp}$ are used in the regression. The resulting number of records for regression of $SA(T = 0.1, 1s)$ are shown in Table 1.

2.2 Parameter selection

- **Moment-magnitude (M):** Laurendeau and Kotha (2025) demonstrated that using the homogenised-ranked M of Laurendeau and Kotha (2025) in place of the ESM-EMEC M reduced the between-event component of the total variability σ of Kotha et al. (2022a) by up to 18% at low frequencies. From this experience, it was deemed necessary to use homogenised M in the regression of compendium flatfile of this study as well. Therefore, the EPOS-ESM M are replaced by the Laurendeau and Kotha (2023) M whenever available. The M of EPOS-FR are mostly from the Si-Hex catalog, but these are replaced with the M estimates from

the regionalised non-parametric spectral decomposition of [Buscetti et al. \(2024\)](#). The M of BELSHAKE dataset were estimated by [Vanneste and Onvani \(2025\)](#) from displacement spectra using the Python package SourceSpec ([Satriano 2022](#)). Trial regressions were performed to evaluate the improvement in σ using different preferred ranking schemes of M , similar to [Laurendeau and Kotha \(2025\)](#). Eventually, the preferred ranking of M followed the order: 1) the spectral decomposition based M for EPOS-FR dataset; 2) the homogenised M of [Laurendeau and Kotha \(2023\)](#) for EPOS-ESM dataset; 3) the default M value provided by the dataset; and 4) the ESM-EMEC M for the EPOS-ESM dataset.

- **Distance (R):** The Joyner-Boore distance ($R = R_{JB}$) is used whenever available. If this is not the case, for events of $M \leq 6.2$, the epicentral distances ($R = R_{epi}$) were used instead. Following several trials, the maximum distance is set to 600km, as further constrains did not effect the model substantially.
- **Site-response (θ):** Most of the stations in the dataset do not have measured V_{s30} . If a site-response parameter were to be included in the GMPM, it is necessary that each site and record is associated with θ - absence of which would make the data unusable in the GMPM regression. Therefore, there is no $G_S(\theta)$ included in the μ of the GMPM, as in $K20$. Instead, site-specific terms are estimated in the mixed-effects regression, which can be subsequently clustered in to site-classes ([Kotha et al. 2018](#)), correlated to a site-response proxy ([Loviknes et al. 2024](#); [Weatherill et al. 2020](#)), spatially smoothed using a geospatial tool ([Parker and Baltay 2022](#); [Chakraborty and Kotha 2026](#); [Chakraborty and Goto 2018](#)) or used as GMPM adjustments in site-specific PSHA ([Rodriguez-Marek et al. 2013](#); [Faccioli et al. 2015](#); [Kotha et al. 2017](#)).

3 Regionalisation models

Figure 1 shows the spatial distribution of records based on event and site locations in the compendium dataset. Unlike the $K20$, in this case, there is a more even distribution of earthquakes across high and low seismicity regions. Nevertheless, Central Italy is by the far the most densely sampled with numerous sites recording more than 3 earthquakes.

3.1 Attenuation regionalisation

The $K20$ attenuation regionalisation models were based on geological and geophysical criteria of [Basili et al. \(2021\)](#); while those in [Weatherill et al. \(2020\)](#) used in ESHM20 are larger *clusters* of $K20$ attenuation regions. This regionalisation enabled $K20$ to quantify the spatial variability of *apparent* anelastic attenuation in the high seismicity Euro-Mediterranean region. However, the $K20$ regionalisation had severe limitations which motivated it's evaluation and revision in this study ([Kotha and Traversa 2024](#)). Please note that from here-on the term *apparent* will be dropped

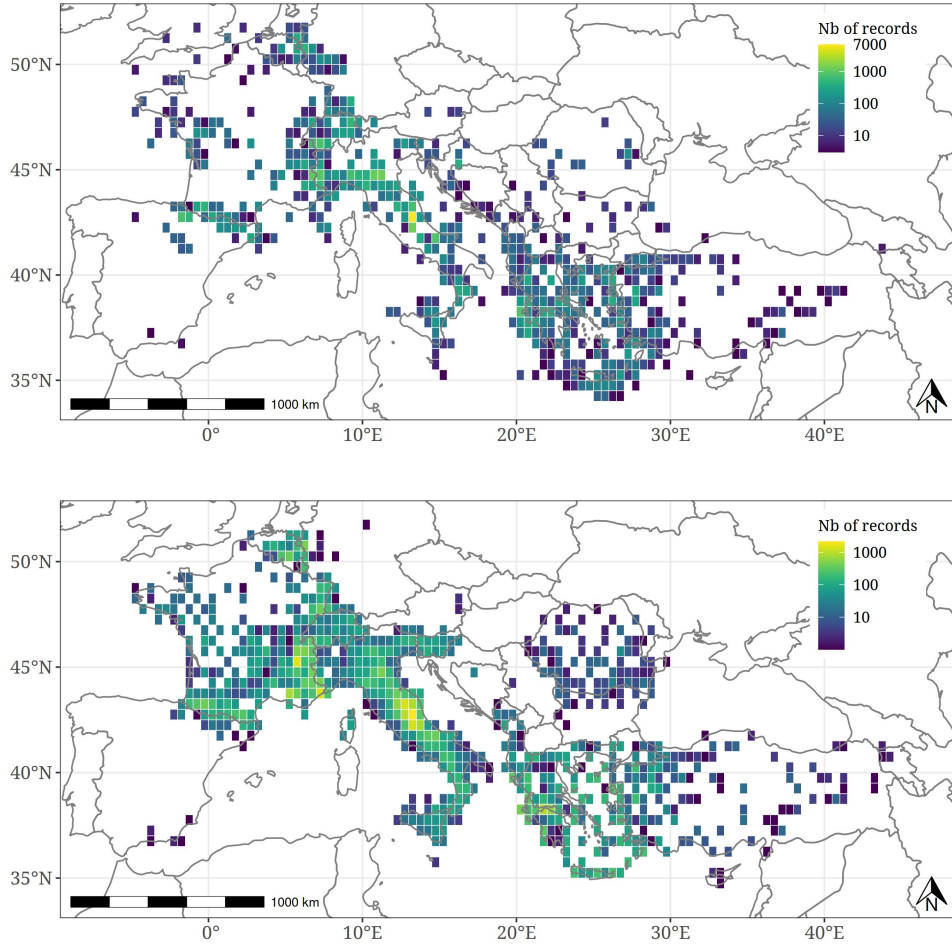


Fig. 1 Geographic distribution of the number of earthquakes (top panel) and stations (bottom panel) in each grid cell. A logarithmic color scale is used to represent the counts and differ for earthquakes and stations.

wherever not necessary.

*K*20 regionalisation of Italy was sufficiently refined to separate regions of low and high attenuation around the Apennines in Central Italy. France, on the other hand, is classified as a single, large region with separate smaller regions covering the Pyrenees and the French-Alps. The attenuation tomography of [Mayor et al. \(2018\)](#) and [Campillo and Plantet \(1991\)](#) revealed as high as 100% variability in coda wave quality factor *Q* for frequencies between 1 and 32 Hz within France. Another demonstration of the importance of including these regional variations in the GMPM is the study of [Pasyanos \(2015\)](#), wherein including 2D variation of crustal shear-wave

attenuation in GMPMs improved their predictions. Alongside, the $K20$ regionalisation was independent of period T (or frequency), while attenuation with distance and its spatial variability are strongly period dependent. In this regard, the $K20$ regional polygons restricted the mixed-effects regression from quantifying statistically significant period-dependent variance of the anelastic attenuation coefficients in $G_P(\mathbf{M}, R)$. Therefore, a more flexible, data-driven regionalisation is sought in this study.

In a few initial trials, the $K20$ regionalisation for France was modified based on ambient-noise velocity tomography of the crust in hexagonal France. Eventually, the most suitable regionalisation model, based on information criteria and analyses of variances, is found to be a regular geographic (lat–lon) grid of $0.5^\circ \times 0.5^\circ$ over the whole of Europe. The grid regionalisation offers a few important advantages:

- The regionalisation is data-driven with no *expert decisions* involved. In $K20$, large regions were assigned regional attenuation adjustments based on a few, sparse ground-motion observations. With the grid regionalisation, absence of data means no regional adjustments will be estimated for that grid cell.
- The attenuation adjustments can vary more gradually in space, instead of sharp transitions across too adjacent regions with very distinct attenuation characteristics.
- In a Bayesian framework, grid-cells with no data can be updated with availability of new data, e.g. using the Uncertainty Project Mapping technique of [Chakraborty and Kotha \(2026\)](#)

3.2 Attenuation sensitivity

In most GMPMs, the anelastic attenuation modelled into their $G_P(\mathbf{M}, R)$ has no exact physical equivalent, but is assumed to be a combination of scattering and absorption of mechanical energy in the propagation medium. Despite being a complex non-linear, non-stationary 3D process, this phenomenon is modelled simplistically in $G_P(\mathbf{M}, R)$ as a stationary, exponential decrease of SA with scalar distance R measures. More recent *partially* non-ergodic GMPMs modelled regional variations in this exponential decrease with R . The more advanced *fully* non-ergodic GMPMs with cell-based attenuation ([Lavrentiadis et al. 2023](#)) were built on the assumption that attenuation is equally sensitive along the entire ray path. However, this assumption may not be appropriate for response spectra, which are influenced not only by the direct seismic phases but also by later-arriving scattered phases. Studies of coda waves have shown that sensitivity to scattering and absorption is not uniformly distributed along the propagation path. Instead, the corresponding sensitivity kernels are primarily concentrated around the source and the recording site. This observation motivates an alternative parametrisation of attenuation, in which the apparent attenuation affecting response spectra is represented by event and site locations rather than by a path-integrated attenuation model. We refer to the studies of [Mayor et al. \(2014\)](#), [Margerin et al. \(2016\)](#), and [Mayor et al. \(2018\)](#) on sensitivity of wave attenuation to perturbations in the medium between source (event) and receiver (site):

- In case of $SA(T)$, the single-degree-of-freedom oscillator transfer functions of period T integrate Fourier amplitudes over a broad frequency range ($f \leq 1/T$) of the body and surface waves, and direct and multiply scattered arrivals.
- [Mayor et al. \(2014\)](#) demonstrated that the attenuation is significantly more sensitive to absorption at the event and site locations, as compared to the wave propagation path. The sensitivity to path increases for shorter event-to-site separation distances. However, the prevalent $G_P(\mathbf{M}, R)$ structure constrains the anelastic attenuation to be significant at larger event-to-site distances. In this sense, distributing the attenuating effect over the entire path may be less important than concentrating them to small *volumes* around event and site locations.
- The regional variability of $G_P(\mathbf{M}, R)$ is calibrated from the numerous small earthquakes, which at sufficiently far-source distances, can be approximated as point sources. However, for larger, finite fault earthquakes, and at near-source distances, scattering in the surrounding fractured media may amplify or attenuate the waves. This means that, as much as the event, site, and path, the attenuation is also sensitive to the scattering volume of the crust.

In addition to the above, and from the lack of sufficient data to constrain fully non-ergodic cell-based attenuation models from ESM18, *K20* modelled a simpler regional attenuation variability based on the site or station location. In the new GMPM, the anelastic attenuation component of $G_P(\mathbf{M}, R)$ is designed to vary with both site locations (as in *K20*) and event locations. The other component of $G_P(\mathbf{M}, R)$ is the geometric spreading, which remains hypocentral depth dependent as in *K20*.

Alongside attenuation regionalisation, *K20* also *localised* earthquakes to quantify systematic differences between ground-motions of events nested in various tectonic localities. The between-locality source variability was meant to capture the source parameter variabilities, contributed to by, for example, regional variability in average stress-drops, biases in $\mathbf{M}$, high-frequency diminution of source spectra ([Kotha et al. 2022b](#); [Laurendeau and Kotha 2025](#)). *K20* used the tectonic localities designed by [Danciu et al. \(2024\)](#) as precursors to seismic area source models of ESHM20. The tectonic locality variability is later modelled as a source related epistemic uncertainty in the ground-motion logic-tree of ESHM20 ([Weatherill et al. 2020](#)). In the new GMPM, with the same arguments made towards revision of the attenuation regionalisation, the *K20* tectonic localities are replaced with a regular grid of $0.5^\circ \times 0.5^\circ$ overlying Europe.

Figure 1 shows the distribution of number of records over the regionalisation (and localisation) grids counted as contributions from events and sites located within each grid cell. The grid-cells with data are those for which event and site location dependent attenuation region and tectonic location dependent adjustments are estimated.

507 4 Parametric form

508 509 4.1 Fixed-effects

510 The parametric mixed-effects form of the GMPM is identical to the one used
 511 in *K20* and explained in [Kotha et al. \(2022b\)](#). The fixed-effects component pre-
 512 dict the natural-logarithm of $SA(T)$: $\mu(T) = G_E(\mathbf{M}, T) + G_P(\mathbf{M}, R, T)$. The
 513 $G_P(\mathbf{M}, R, T)$ consists of the geometric spreading $G_{P,g}(\mathbf{M}, R, T)$ and anelastic atten-
 514 uation $G_{P,a}(\mathbf{M}, R, T)$ components. The regressions are performed at each T inde-
 515 pendently, i.e. without assuming cross-correlation of $SA(T)$ and the fixed-effects
 516 coefficients; therefore, T is not shown in the equations subsequently. The $SA(T)$ spatial
 517 cross-correlation models for this GMPM are provided by ([Damiani et al. 2026a,b](#)).

$$519 \quad G_E(\mathbf{M}) = e_1 + \begin{cases} b_1(\mathbf{M} - M_h) + b_2(\mathbf{M} - M_h)^2 & \mathbf{M} \leq M_h \\ b_3(\mathbf{M} - M_h) & M_h < \mathbf{M} \end{cases} \quad (1)$$

522 In the $\mathbf{M}$ -scaling component of the GMPM shown in Equation (1), $M_h = 5.7$ is
 523 the T -independent *hinge*-Magnitude in [Kotha et al. \(2022b\)](#). For $\mathbf{M} \leq M_h$, $SA(T)$
 524 scale quadratically with $\mathbf{M}$; and linearly with $\mathbf{M}$ for $M_h < \mathbf{M}$.

$$526 \quad G_{P,g}(\mathbf{M}, R) = (c_1 + c_2(\mathbf{M} - M_{ref})) \ln \sqrt{\frac{R^2 + h_D^2}{R_{ref}^2 + h_D^2}} \quad (2)$$

$$529 \quad G_{P,a} = \frac{c_3}{100} (\sqrt{R^2 + h_D^2} - \sqrt{R_{ref}^2 + h_D^2}) \quad (3)$$

530 In the geometric spreading component shown in Equation (2) and the anelas-
 531 tic attenuation component shown in Equation (3), as in *K20*, $R_{ref} = 30km$ is the
 532 T -independent reference distance, and h_D is the h-parameter or the pseudo-depth
 533 parameter required to emulate near-source saturation of $SA(T)$. h_D is set as a prior
 534 based on the hypocentral depth (D) as follows:

$$537 \quad h_D = \begin{cases} 4 & D \leq 10km \\ 8 & 10km < D \leq 20km \\ 12 & 20km < D \end{cases} \quad (4)$$

540 The coefficients $e_1, b_1, b_2, b_3, c_1, c_2, c_3$ constitute the fixed-effects of the GMPM and
 541 are *generic* to the entire dataset, i.e. they are not *specific* to any particular event, site,
 542 attenuation region or tectonic locality.

544 4.2 Random-effects

546 Random-effects are adjustments to the fixed-effects coefficients required to make pre-
 547 dictions specific to any event, site, attenuation region or tectonic locality. As with the
 548 fixed-effects coefficients, random-effects are also estimated independently for each T .
 549 The random-effects structure is shown in Equation (5) below:

$$551 \quad (e_1|\delta B_{e,l}^0) + (e_1|\delta S2S_s) + (e_1|\delta L2L_l) + (c_3|\delta c_{3,r}^e) + (c_3|\delta c_{3,r}^s) \quad (5)$$

- $\delta B_{e,l}^0$ is the between-event random-effect or term of event e nested in the tectonic locality l . $\delta B_{e,l}^0$ are sampled from the random distribution $\Delta B^0 = \mathcal{N}(0, \tau_0^2)$, where τ_0 is the between-event variability. $\delta B_{e,l}^0$ are event-specific adjustments to the fixed-effects coefficient e_1 . Since $\delta B_{e,l}^0$ of a prospective event cannot be predicted, τ_0 is considered part of the GMPM's aleatory variability σ . 553
554
555
556
557
- $\delta S2S_s$ is the between-site term of site s . $\delta S2S_s$ are sampled from the random distribution $\Delta S2S = \mathcal{N}(0, \phi_{S2S}^2)$, where ϕ_{S2S} is the between-site variability. $\delta S2S_s$ are site-specific adjustments to the fixed-effects coefficient e_1 . Assuming the site-response remains linear and stationary, $\delta S2S_s$ can be considered as repeatable site-effects, and used in site-specific PSHAs. For a site without $\delta S2S_s$, however, site-response proxies maybe used predict a partially ergodic $\delta S2S_s$ or include ϕ_{S2S} in the σ . 558
559
560
561
562
563
564
- $\delta L2L_l$ is the between-locality term of tectonic locality l . Unlike in *K20*, where tectonic localities l were large polygons, in this GMPM the l refer to the grid-cells shown in Figure 1. $\delta L2L_l$ are sampled from the random distribution $\Delta L2L = \mathcal{N}(0, \tau_{L2L}^2)$, where τ_{L2L} is the between-locality variability. $\delta L2L_l$ are locality-specific adjustments to the fixed-effects coefficient e_1 . $\delta L2L_l$ is roughly equal to the mean of the *non-localised* δB_e of the events nested in the locality l , as in : $\delta B_e = \delta L2L_l + \delta B_{e,l}^0$. $\delta L2L_l$ are estimated from several small-moderate sized earthquakes in a locality l , and are assumed to be repeatable (stationary) effects. Therefore, *K20* proposed using $\delta L2L_l$ as locality-specific adjustments to e_1 . In case the $\delta L2L_l$ have large uncertainties, instead of including τ_{L2L} in the σ , it can be treated as an epistemic uncertainty in ground-motion logic-trees. For example, ESHM20 ground-motion logic-tree developed by Weatherill et al. (2024) considers τ_{L2L} as a source stress-parameter uncertainty. In this study, τ_{L2L} is not treated as an aleatory variability component but preferably used in a ground-motion logic-tree as an epistemic uncertainty. 565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
- $\delta c_{3,r}^e$ is the adjustment made to anelastic coefficient c_3 of $G_{P,a}$ (Equation (3)), depending on the region r containing the event e . In this case, the region r are grid-cells, unlike the regional polygons used in *K20*. Similarly, $\delta c_{3,r}^s$ are adjustments based on the grid-cell containing the site s . The $\delta c_{3,r}^e$ and $\delta c_{3,r}^s$ attenuation random-effects are sampled from the distributions $\Delta c_3^e = \mathcal{N}(0, \tau_{c3,e}^2)$ and $\Delta c_3^s = \mathcal{N}(0, \tau_{c3,s}^2)$, respectively. As with τ_{L2L} , $\tau_{c3,e}$ and $\tau_{c3,s}$ can be used as epistemic uncertainties in a ground-motion logic-tree similar to that used in ESHM20. Alternatively, if the regions containing the event and site have an associated $\delta c_{3,r}^e$ and $\delta c_{3,r}^s$, they can be used to estimate an event-and-site dependent adjustment to anelastic attenuation coefficient $c_3 + \delta c_{3,r}^e + \delta c_{3,r}^s$ in $G_{P,a}$. In this study, $\tau_{c3,e}$ and $\tau_{c3,s}$ will be treated as epistemic uncertainties on attenuation, rather than as a component of the GMPM's aleatory variability. 580
581
582
583
584
585
586
587
588
589
590
591
592

The random-effects structure shown in the Equation (5) is verified to be the most efficient in terms of the GMPMs explanatory (e.g. analyses of variances), and prediction power (e.g. k-fold cross validation) in *K20*, and discussed in section 7. It is to be noted that all the above random-effects are assumed to follow a Gaussian distribution ($\mathcal{N}$) because the regression algorithm assumes the natural-logarithm of ground-motion observations are sampled from a Gaussian distribution; and therefore, 593
594
595
596
597
598

the random-effects distributions would be Gaussian as well. Random-effects group variances (e.g. τ_0^2) are estimated first, from which random-effects values (i.e. $\delta B_{e,l}^0$) are sampled for each level (e.g. event e). However, normality tests (Massey Jr. 1951) often reject the null-hypothesis that random-effects are sampled from a symmetric Gaussian distribution. Moreover, random-effects are not residuals, but are adjustments to fixed-effects coefficients in a mixed-effects GMPM. In this GMPM, $\delta L2L_l$, $\delta c_{3,r}^e$, and $\delta c_{3,r}^s$ are considered repeatable effects, and their standard-deviations τ_{L2L} , τ_{c3}^e and τ_{c3}^s , respectively, are not included in the σ of the GMPM. Instead, τ_{L2L} , τ_{c3}^e and τ_{c3}^s can be used as epistemic uncertainties in ground-motion logic-trees; or use $\delta L2L_l$, $\delta c_{3,r}^e$, and $\delta c_{3,r}^s$ as fixed-effects adjustments can be to make tectonic locality and attenuation region specific predictions based on event e and site s locations.

The *apparent* natural randomness of the ground-motion observations is contained in the so-called *left-over* residuals $\epsilon_{e,s}$, indexed by event e recorded at site s . These residuals are assumed to follow a Gaussian distributions $\mathcal{N}(0, \phi_0^2)$ as well. ϕ_0 is the record-to-record variability of the GMPM, and may contain secondary phenomena that are unrepeatable (not random-effects), unpredictable (not fixed-effects) or both. ϕ_0 together with the between-event aleatory variability τ_0 constitute the aleatory variability of the GMPM, excluding site-response variability. $\delta S2S_s$ is a more well-known repeatable site-effects, and can be used in site-specific predictions, mapped over large regions either using geostatistical tools (Parker and Baltay 2022; Bergamo et al. 2025; Chakraborty and Kotha 2026) or correlated to site-response proxies (Weatherill et al. 2023; Loviknes et al. 2024; Lanzano et al. 2020). In doing so, the total ϕ_{S2S} need not be included in the σ of the GMPM.

4.3 Regression algorithm

The improvement in seismic networks, automatic processing of waveforms (Thompson et al. 2025; Zaccarelli et al. 2019), and data retrieval web services (Mascandola and Sgobba 2025) have facilitated generation of very large, high-quality ground-motion datasets on-demand. Alongside, GMPMs have become increasingly complex, more dataset and application specific (Bindi et al. 2017), and numerous (Douglas 2024). Consequently, GMPM selection in ground-motion logic-trees (Cotton et al. 2006) has become more challenging considering the fact that many GMPMs are regressed from nearly identical datasets and tend to predict either very similar or very dissimilar ground-motion distributions (Douglas 2024). Another less acknowledged aspect is that most GMPMs do not report the complete set of their epistemic and aleatory uncertainties (Liou and Abrahamson 2024), which could be vital in model selection. This means that, with increasing amount of data and the increasing number of alternative GMPMs, both epistemic and aleatory uncertainties in ground-motion prediction have increased. One of the solutions to this issue is to develop GMPMs in a Bayesian framework; where-in, the within-model uncertainties are quantified, investigated, and reduced with the availability of new data (Stafford 2019; Kowsari et al. 2019; Kuehn 2021). To this end, the regionally adaptable *backbone* GMPM in this study is derived using the Integrated Nested Laplacian Approximation of Bayesian

regressions (INLA, Gómez-Rubio (2020)).

Bayesian linear mixed-effects regressions, unlike the frequentist approach used in *K20*, estimate the full posterior distribution of all mixed-effects of a GMPM. The posterior distributions provide a complete picture of the GMPM's uncertainties, and more importantly, act as informative priors in updating selected mixed-effects components with new data from new regions (Kotha and Traversa 2024). However, this is possible only when the mixed-effects structure of the GMPM does not need to be changed to suit the new dataset, and if the metadata (e.g. $\mathbf{M}$) are not revised. In this study, the attenuation regionalisation and tectonic localities have changed, an additional random-effect is introduced, and a substantial number of $\mathbf{M}$ are revised with respect to *K20*. Therefore, the mixed-effects of *K20* are not used as information priors, and a new GMPM is regressed from scratch using the Bayesian INLA algorithm (R-INLA, Lindgren and Rue (2015)).

5 Results

5.1 Fixed-effects

The fixed-effects component of the GMPM predicts the natural-logarithm of $SA(T)$: $\mu(T) = G_E(\mathbf{M}, T) + G_P(\mathbf{M}, R, T)$. Figure 2 shows the 5% damped RotD50 (gal, cm.s^{-2}) measure of $SA(T = 0.1\text{s})$ (top-panels) and $SA(T = 1\text{s})$ (bottom-panels). The plots compare the R -scaling for $\mathbf{M} = 4, 5.5, 7$ (left-panels), and the $\mathbf{M}$ – scaling for $R = 10, 50, 150\text{km}$ (right-panels) of *K20* (yellow curves) against the new GMPM (black curves). Note that the fixed-effects coefficients of *K20* were updated in Kotha et al. (2022b), and the plots shown here are for *Intermediate* hypocentral depths $10\text{km} < D \leq 20\text{km}$ and *Average* attenuation versions of *K20*. The epistemic uncertainty on fixed-effects prediction ($\pm 1\sigma_\mu$ discussed subsequently) is presented as an uncertainty band around the median. This allows confirming if the difference in fixed-effects predictions from the two models are statistically significant. Figure 3 compares the fixed-effects coefficients and the random-effects standard-deviations of the new model (solid lines) against those of the *K20* (dashed lines). This plot allows interpreting the comparison in Figure 2.

The R -scaling (left-panels) suggests that the new GMPM (black) predicts lower $SA(T = 0.1\text{s})$ (top-panels) and $SA(T = 1\text{s})$ (bottom-panels) than *K20* (yellow) for $\mathbf{M} \leq 7$ at $R \leq 80\text{km}$ distances. At $R \geq 80\text{km}$ distances the two models are practically identical, given the marginal differences between predicted weak ground-motions ($\leq 1 \text{ cm.s}^{-2}$). Only for $\mathbf{M} \geq 7$, the ground-motion are significantly higher as interpreted from the non-overlap of $\pm 1\sigma_\mu$ uncertainty bands. Referring to the c_1, c_2, c_3 coefficients of $G_P(\mathbf{M}, R, T)$ in Figure 3, the estimates from the two models are almost identical. Despite the 3-fold increase in data from more diverse regions than in *K20*, inclusion of an additional random-effect, a revised regionalisation model, and revision of several $\mathbf{M} \leq 5.5$ values, the $G_P(\mathbf{M}, R, T)$ is practically unaffected. This means that, the $G_P(\mathbf{M}, R, T)$ is uncorrelated to other mixed-effects of the GMPM,

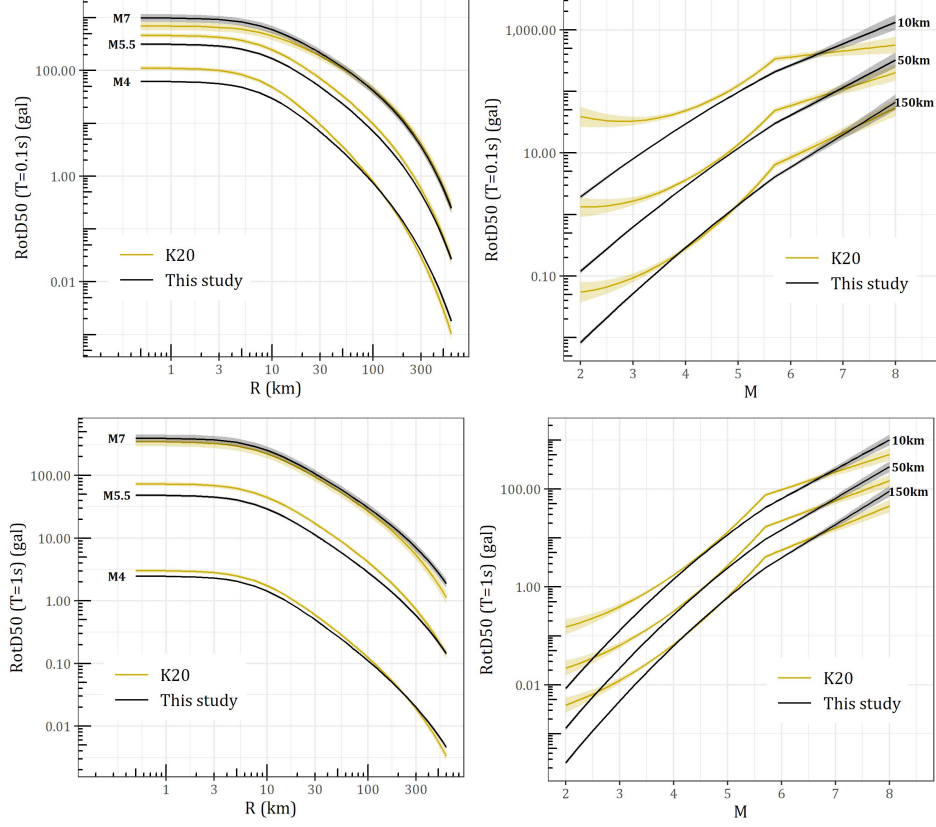


Fig. 2 Parametric representation of the spectral acceleration (SA) predictions: distance scaling (left panels) for four earthquake magnitudes (M 3, 4, 5.5, and 7.5), and magnitude scaling (right panels) for three source-to-station distances (10, 50, and 150 km). Top panels show SA at $T = 0.1$ s, bottom panels at $T = 1.0$ s. Black curves represent the present model; yellow curves represent the model of [Kotha et al. \(2022b\)](#).

which was also demonstrated in [Kotha et al. \(2022b\)](#). The change in predictions at near-source distances for $M > 6.5$ is from the change in the $M - scaling$.

The $M - scaling$ (right-panels) reveal striking differences between $K20$ and the new GMPM. The new GMPM is regressed over a wider $1.6 < M \leq 7.7$ range than the $3.1 < M \leq 7.4$ of $K20$. As a result, the new GMPM has greater range of applicability than $K20$. The most remarkable change is the behavior at $M \leq 5$. As discussed [Kotha et al. \(2022b\)](#), the $M - scaling$ of $K20$ was rather *nonphysical*, because it predicted a quadratic increase in ground-motions towards $M < 4.5$. This was inconsistent with the theoretical expectation that the smaller earthquakes are likely to have lower radiation efficiency, and therefore, should produce ground-motions much lower than a linear decay with decreasing M ([Douglas and Aochi 2016](#)). To imitate this physics, the coefficient b_2 of $G_E(M, T)$, controlling the quadratic scaling of $SA(T)$ with M , is supposed to be a negative value ([Fukushima 1996](#)). However,

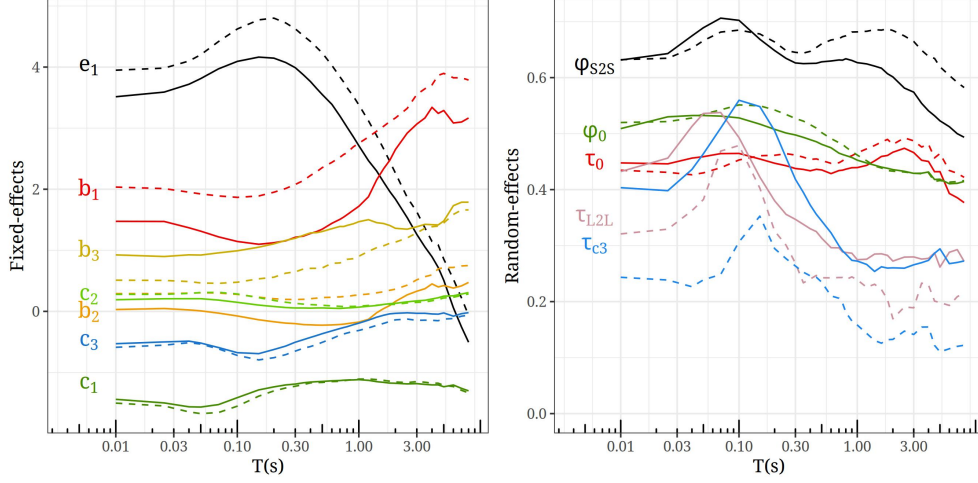


Fig. 3 Frequency dependence of fixed-effects coefficients and random-effects standard-deviations compared to the 2022 update of the K20 (Kotha et al. 2022b). Solid lines represent the present model; dashed lines represent the model of Kotha et al. (2022b).

as seen in Figure 3, b_2 of K20 is positive for all T . Kotha et al. (2022b) argued that this unusual behavior is due to the fact that most of $3.5 < M \leq 4.5$ values in the regressed dataset were inferred from regionally calibrated M_L estimates; making them possibly erroneous. It was, however, decided against removing $3.5 < M \leq 4.5$ data from the regression of K20 because the numerous small-moderate sized earthquakes were necessary to constrain the random-effects; and the applicability of K20 was strictly limited to $4.5 < M$. Thanks to the efforts of Vanneste and Onvani (2025), Buscetti et al. (2024), and Laurendeau and Kotha (2025) in providing more reliable, homogenised estimates of $1.6 < M \leq 5$, the new GMPM behaves more physically meaningful, with its $b_2 \leq 0$ for $T \leq 3s$. The new GMPM can now be used for predictions of $M \leq 4.5$ with greater confidence.

Another major difference in M -scaling of the two GMPMs is around and above the $M_h = 5.7$ - the hinge magnitude. The K20 GMPM predicts larger $SA(T)$ at $5.7 < M \leq 7$, and exhibits a *sharp* break at $M_h = 5.7$. The new GMPM shows a more realistic smoother transition between moderate and large M earthquakes separated by $M_h = 5.7$. Fixed-effects coefficient b_3 guides the M - *scaling* beyond $M_h = 5.7$, and is more positive in the new GMPM than in K20 (Figure 3). This behavior is better evaluated in the analyses of the within-model epistemic uncertainty (σ_μ) and $\delta B_{e,l}^0$ random-effects in subsequent sections.

The electronic supplement contains a plot comparing the 5% damped RotD50 response spectra of the two models (Figure S1). The two models have similar shapes, with increase in *predominant* period with increasing M , imitating the decrease in corner-frequency (f_c) of ω^{-2} source spectrum (Brune 1970) with increasing M . The

differences in the response spectra of the two models are more pronounced at $T \geq 1/f_c$ at $R \leq 30\text{km}$, and become negligible at larger distances. All of these differences can be explained from the comparisons in Figure 2 and Figure 3.

5.1.1 Epistemic uncertainty on fixed-effects

The within-model epistemic uncertainty σ_μ is an $[\mathbf{M}, R, T]$ dependent standard-deviation computed from the fixed-effects coefficients' covariance matrices and the gradients of median μ with respect to each fixed-effects coefficient (Atik and Youngs 2014). $\sigma_\mu(\mathbf{M}, R, T)$ is a useful measure to evaluate if the complexity of GMPM's fixed-effects is justified in terms of the sufficiency of data to constrain their fixed-effects coefficients. The $\sigma_\mu(\mathbf{M}, R, T)$ of the new GMPM is compared to that of $K20$ reported in Kotha et al. (2022b) in Figure 4. The left column in this plot corresponds to the new GMPM, the central column corresponds to $K20$, and the right column shows the difference $\Delta\sigma_\mu(\text{This study} - K20)$, for $T = 0.1\text{s}$ in top and $T = 1\text{s}$ in bottom rows. It is evident that the new GMPM has substantially smaller $\sigma_\mu(\mathbf{M}, R, T)$ than $K20$. The new GMPM has practically no epistemic uncertainty at $\mathbf{M} \leq 6.5$. At $\mathbf{M} > 6.5$ the epistemic is only slightly lower than that of $K20$ because there is not much new data since the Amatrice-Visio-Norcia 2016 sequence in Central Italy, which was already included in the $K20$ regression. Towards longer periods, the new GMPM has even lower epistemic uncertainty at $\mathbf{M} > 6.5$ than the $K20$, which further supports its replacement of $K20$.

5.2 Random-effects

5.2.1 Random-effects and residual standard-deviations

The right panel of Figure 3 shows the random-effects and residual and standard-deviations of the GMPM (solid lines) relative to those of $K20$ (dashed lines). σ is the total ergodic aleatory variability of a GMPM. Depending on the application, however, non-ergodic variants of σ can be used by dropping certain random-effects standard-deviations from its estimation. For example, for site-specific PSHA, the between-site standard-deviation ϕ_{S2S} can be removed from σ , provided the $\delta S2S_s$ of the site is available. Therefore, to avoid cluttering, Figure 3 does not compare the ergodic or non-ergodic σ of the two GMPMs.

Here it is important to mention that, while the published version of $K20$ was derived using robust linear mixed-effects regression, the standard-deviations compared here are from a version using Bayesian INLA algorithm (Kotha and Traversa 2024). At the time of this study, the INLA algorithm was not capable of robust estimates of random-effects and residual standard-deviations. Nevertheless, the robust counterparts of the random-effects and residual standard-deviations can be estimated as the Interquartile Ranges (IQR) of their distributions. The IQR quantifies the statistical dispersion of the central half of a random distribution. Unlike the Gaussian standard-deviation containing middle 68% of a distribution, the IQR focuses on the middle 50%, making it less sensitive to outliers and extreme values. For practical

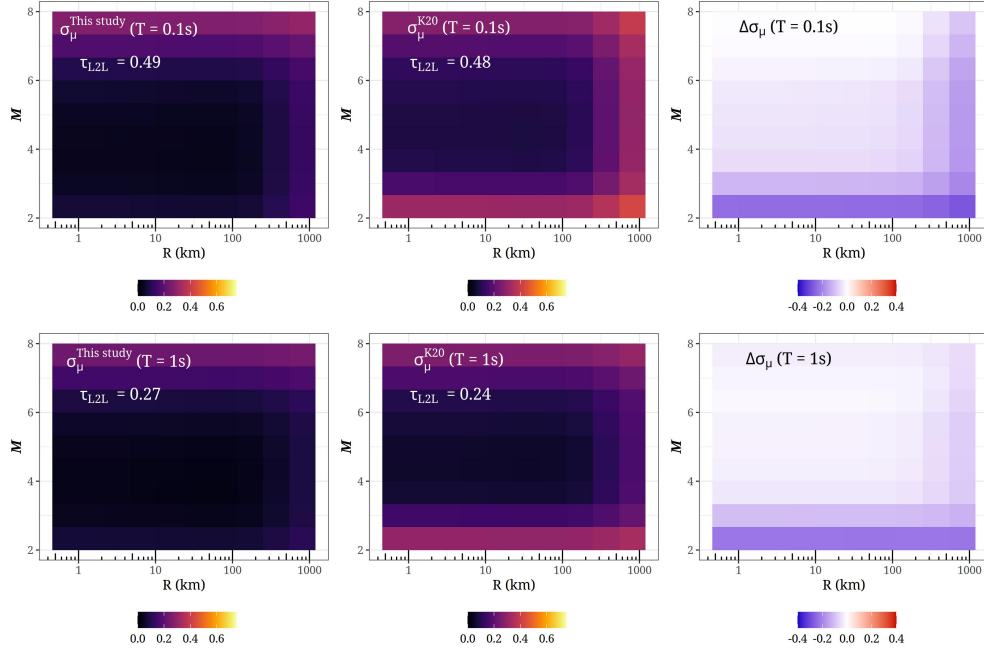


Fig. 4 $\sigma_\mu(M, R, T)$ of the new GMPM compared to *K20* model reported in [Kotha et al. \(2022b\)](#). The left columns present new GMPM, the central column corresponds to *K20*, and the right column shows the difference $\Delta\sigma_\mu(\text{This study} - \text{K20})$. The upper figures shows for $T = 0.1s$ and the lower ones for $T = 1s$.

application of the new GMPM, e.g. in PSHA, the robust variants of random-effects and residual standard-deviations are estimated as $0.7413 \cdot \text{IQR}$.

Table 1 shows the number of records, sites, events, site attenuation regions, event attenuation regions, and tectonic localities quantifying the estimates of ϕ_0 , ϕ_{S2S} , τ_0 , $\tau_{c3,s}$, $\tau_{c3,e}$, τ_{L2L} in Figure 3 at $T = 0.1$ and $1s$. Cross-referencing Table 1 and Figure 3, it is evident that the enormous increasing in number of records and events has not significantly impacted the ϕ_0 . τ_0 at $T < 0.1s$ has increased compared to that of *K20*, which is to be expected with the 3-fold increase in number of events in the new dataset, most of which are of $M < 4.5$. At $T \geq 0.1s$, τ_0 has decreased with respect to *K20*, which is likely from the homogenised M of small-moderate sized earthquakes. The reduction in both ϕ_0 and τ_0 at $T \geq 0.1s$ would reflect as reduction in the non-ergodic $\sigma_0 = \sqrt{\tau_0^2 + \phi_0^2}$ of the new GMPM compared to *K20*.

The remaining random-effects variances need to be included in the σ of the GMPM, but instead used in non-ergodic predictions or as epistemic uncertainties in GMPM logic-trees. The number of sites has doubled since *K20*, yet there is a decrease in ϕ_{S2S} at $T \geq 0.3s$. The number of tectonic localities has increased 10-fold in terms of both the spatial extent and resolution, which led to a steep

increase in τ_{L2L} . The same applies to quantified attenuation variability τ_{c3} . Here, since the *K20* only has the site location dependent τ_{c3} , the new GMPM's τ_{c3} is estimated as $\sqrt{\tau_{c3,e}^2 + \tau_{c3,s}^2 + 2\text{Cov}(\Delta c_3^e, \Delta c_3^s)}$, where Δc_3^e and Δc_3^s are the event and site dependent attenuation random-effects distributions. It is worth reiterating that, the notations with Δ indicate the random-effects *distributions*, while those with δ are *specific values* from the distribution. The increase in τ_{L2L} and τ_{c3} need not be treated as increase in aleatory variability and a larger ergodic σ , because they can be used as epistemic uncertainties in a ground-motion logic-tree, as done in ESHM20 (Weatherill et al. 2020). In summary, the fixed-effects epistemic uncertainty and random-effects aleatory variability are both lower for the newer model, but there is an increase in random-effects epistemic uncertainty for attenuation regions and tectonic localities not featuring in the regressed data. Nevertheless, as the new model is cast in a Bayesian framework, it can be updated with availability of new data.

5.2.2 Anelastic attenuation variability

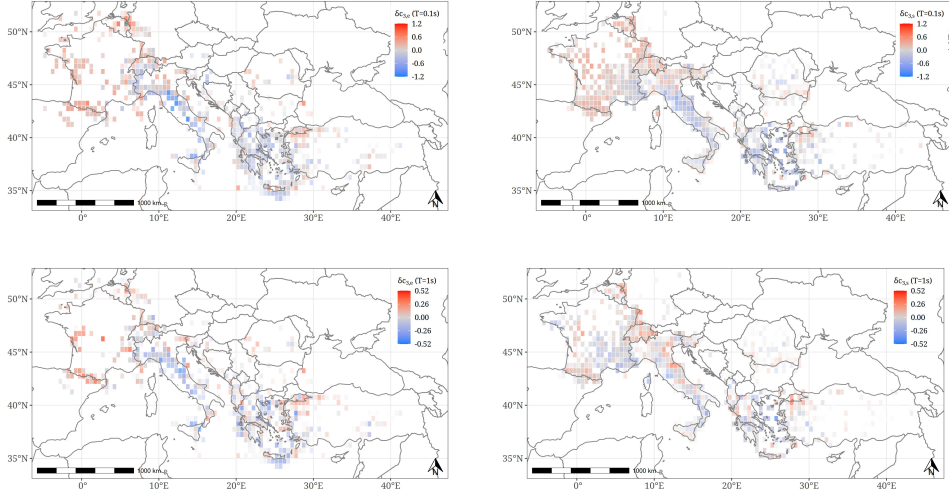


Fig. 5 Spatial distribution of anelastic attenuation variability across Europe. Upper panels: $T = 0.1$ s; lower panels: $T = 1.0$ s. Left: source-based attenuation (colored for cells containing earthquakes); right: site-based attenuation (colored for cells containing stations). Blue colors indicate regions with lower ground motion due to higher attenuation relative to the median prediction; red colors indicate regions with higher ground motion due to lower attenuation relative to the median prediction.

Figure 5 shows the spatial variability of apparent anelastic attenuation according to the event (Δc_3^e in Figure 5, left panels) and site locations (Δc_3^s in Figure 5, right panels) for $T = 0.1$ s (upper panels) and $T = 1$ s (lower panels). Regions in red are those with weaker than average attenuation (δc_3^e and $\delta c_3^s > 0$), those in blue are regions of stronger than average attenuation (δc_3^e and $\delta c_3^s < 0$), and those in grey are close to the global average. The less saturated regions (or grid-cells) are those with

larger epistemic uncertainty either from insufficient data to constrain their δc_3^e and δc_3^s or large within-region variability despite sufficient data. The standard-errors on δc_3^e and δc_3^s can be used in a GMPM logic-tree when performing a region-specific partially non-ergodic PSHA (Kotha et al. 2017).

The spatial variability at $T = 0.1s$ shows greater contrast, indicating that attenuation of short-period (or high-frequency) ground-motions is more variable than at $T = 1s$ and longer periods. This can also be inferred from the *peaking* of τ_{c3} in Figure 3 at $T = 0.1s$. It is also worth noting that, despite being the best sampled regions, the regions in Central Italy have non-zero δc_3^s and δc_3^e . This means that, the random-effects distributions are unlikely to be biased by the best sampled regions; which indicates that the generic c_3 fixed-effects coefficient is unbiased as well.

Qualitatively, the spatial variability in the $T = 0.1s$ maps (in Figure 5) is coherent with the larger regional polygons of $K20$. However, unlike in $K20$, regions without sufficient data are assigned δc_3^e and $\delta c_3^s = 0$. In addition, δc_3 values are not assigned over very large regions despite having only a couple of isolated sites within the region. An illustrative example is the improvement in representation of attenuation variability within France. In $K20$, France was assigned $\delta c_3 = 0$ because there was not data available in most regions. In Kotha and Traversa (2024), where $K20$ was updated using the Traversa et al. (2020), the large regional polygon covering most of France was assigned a large $\delta c_3 > 0$. While the new regionalisation shows more refined spatial variability within France. The Pyrenees at the France-Spain frontier are characterised as low attenuation regions ($\delta c_3 > 0$) and distinct from the high attenuation regions of western French-Alps at the France-Switzerland-Italy frontier ($\delta c_3 < 0$). Similar is the case with Belgium, north-western Turkey, southern Romania, and northern Italy; where greater spatial variability is noticeable. Central Italy and Greece are characterised as those with strongest attenuation, while the lower Rhine Graben at the frontier of France-Switzerland-Germany show weakest attenuation; which is similar to $K20$. The spatial variability of δc_3^e and δc_3^s for $T = 0.1s$ is also coherent with intrinsic absorption Q maps of Campillo and Plantet (1991), Mayor et al. (2018) and Heller et al. (2025) for $f = 10Hz$. Although not emphasized in these maps, the spatial patterns also gradually change from short to long-periods, which is to be expected of frequency-dependent anelastic attenuation. However, a thorough investigation of frequency-dependencies is more appropriate in the Fourier domain using the estimates from a Fourier mixed-effects GMPM (Kotha et al. 2022a).

Finally, as most events and sites are not collocated, the δc_3^e and δc_3^s maps are spatially complimentary. For grid-cells containing both events and sites, the correlation between δc_3^e and δc_3^s is not particularly strong. This could be due to two reasons: 1) statistically, part of the attenuation effect could be absorbed into the other event ($\delta B_{e,l}^0$, $\delta L2L_l$), and also likely the site dependent ($\delta S2S_s$) random-effect; 2) physically, while δc_3^s could be quantifying attenuation in the shallow layers at surface sites, the δc_3^e could be quantifying attenuation at a variety of hypocentral depths up to 40km.

5.2.3 Tectonic locality variability

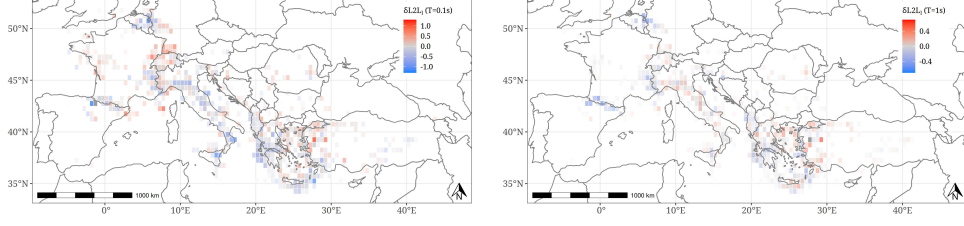


Fig. 6 Locality-to-locality variability ($\Delta L2L$) across Europe at $T = 0.1$ s (left panel) and $T = 1$ s (right panel). Values are shown for each cell of the regionalisation. Blue regions indicate localities where earthquakes produce higher ground motion than the median model prediction; red regions indicate localities where earthquakes produce lower ground motion than the median model prediction.

The tectonic locality-to-locality spatial variability ($\Delta L2L$) is shown in Figure 6, for $T = 0.1$ s (left panel) and $T = 1$ s (right panel). This random-effect captures the systematic deviation in ground-motions of events within a grid-cell (or locality) with respect to the fixed-effects predictions. The τ_{L2L} in Figure 3 also has a peak value at around $T = 0.1$ s - even higher than τ_0 - but rapidly decreases towards longer periods. The left panel showing spatial variability of $\Delta L2L(T = 0.1$ s) shows more contrast (in color) than that of $\Delta L2L(T = 1$ s) in the right panel of Figure 6. This indicates that the short-period ground-motions are more variable across tectonic localities than long-period ground-motions, and also that the spatial patterns are period or frequency dependent.

Referring to the $\Delta L2L(T = 0.1$ s) maps, blue localities are those where the events have produced weaker than average short-period ground-motions. The Pyrenees, the eastern Alps, the volcanic arc in Greece, Central Italy, and Belgium are examples of those regions that appear to have produced weaker short-period ground-motions. Coincidentally, these are the more active tectonic localities characterised mainly by fractured crust and/or ancient mountain ranges. Conversely, the localities in red, e.g. the Swiss basin, the Po Plain, and central Greece, are generally sedimentary basins. In ESHM20, τ_{L2L} is assumed to have captured the regional variability of event stress-drops and modeled as such in the ground-motion logic-tree (Weatherill et al. 2024), but Kotha et al. (2022a) reported that the $\delta_{L2L}(T = 0.1$ s) are better correlated to the events' high-frequency spectral decay parameter κ_{source} - a best-fit *slope* quantifying the decay of high-frequencies (≥ 10 Hz) of the Brune's source spectra estimated by Bindi and Kotha (2020). To perform a similar comparison for this dataset would require the source parameter estimation (as in Bindi and Kotha (2020)) and an FAS GMPM (as in Kotha et al. (2022a)).

5.2.4 Event variability

The between-event random-effect (ΔB^0) captures a part of the *true* aleatory variability of ground-motions, but may also be inflated by uncertainties or errors in the $\mathbf{M}$ estimates of the events. Several studies have demonstrated that the short-period and high-frequency ΔB^0 are strongly correlated to the events' stress-drops, while the long-period and low-frequency ΔB^0 are more correlated to Brune's quality factor Q_0 . A few studies (Laurendeau and Kotha 2025; Kuehn and Abrahamson 2017) demonstrated that the long-period or low-frequency τ_0 can be reduced with improved, homogeneous $\mathbf{M}$. Neither stress-drop nor quality factor are predictable within current state of knowledge, while their estimation itself have several trade-offs and uncertainties (Cochran et al. 2024). In this regard, τ_0 has to be included in the aleatory σ of the GMPM. τ_0 of this and K20 GMPMs are smaller than the traditional between-event random-effect ($\Delta B = \mathcal{N}(0, \tau^2)$) standard-deviation τ , because the repeatable tectonic locality-specific event characteristics are quantified in to τ_{L2L} and used as epistemic uncertainty in ground-motion logic-trees, as in : $\tau_0 = \sqrt{\tau^2 - \tau_{L2L}^2}$.

Figure 7 presents the typical ΔB^0 versus $\mathbf{M}$ diagnostic plot for $T = 0.1s$ in the left and $T = 1s$ in the right panels. The dashed red lines mark $\pm\tau_0$, the red dots mark the bin-wise median values of ΔB^0 . The red error-bars indicate the bin-wise $\pm 0.7413 \cdot \text{IQR}$. The grey scale shows the number (or density) of events. For example, the largest number of events are concentrated around $\delta B_{e,l}^0$ and $4 \leq M < 5$. In both panels, the bin-wise median values of ΔB^0 align with the zero baseline; this means that, there is negligible bin-wise bias in the $\mathbf{M}$ -scaling of the GMPM. However, in the left-panel corresponding to $T = 0.1s$, the error-bars are *wider* towards smaller $\mathbf{M}$ and become gradually *narrower* towards larger $\mathbf{M}$. This indicates a M -dependent heteroscedasticity of τ_0 , also called heterogeneity of variance. On the contrary, at $T = 1s$, the heteroscedastic trend shows maximum variability in the $4 \leq M < 5$ bin, and decreasing towards higher and lower M . To verify the hypothesis of M -dependent heteroscedasticity of τ_0 , a Studentised Breusch-Pagan (BP) test is performed on all periods (Breusch and Pagan 1979). The BP test assesses whether the variance of the residuals depends on the values of the independent variables, in this case ΔB^0 are the residuals and $\mathbf{M}$ is the independent variable. The result of this test (Figure S2 in the supplementary materials), shows that the studentised BP statistic is well above the reference value, suggesting that the null-hypothesis of absence of heteroscedasticity can be rejected at 95% significance.

The BP test confirms the hypothesis of M -dependent heteroscedasticity of ΔB^0 , and therefore, supports a M -dependent τ_0 . The K20, Weatherill et al. (2024), Al Atik (2015), and other GMPMs fit a piece-wise linear $\tau_0(\mathbf{M})$ over bin-wise standard-deviations. For this GMPM, we fit a inverse-hyperbolic trigonometric function $\text{arcsinh}(x) = \ln(x + \sqrt{x^2 + 1})$ to the bin-wise $0.7413 \cdot \text{IQR}$ values. The $\text{arcsinh}(\mathbf{M} - M_{ref})$ in Equation (6) ensures that the principal value $\tau_0(M_{ref})$ of the arcsinh multi-function is fixed at $0.7413 \cdot \text{IQR}$ of the best sampled bin containing $(M_{ref}) = 4.5$, and $\tau_0(\mathbf{M})$ values elsewhere in the domain of definition are defined to agree with those found by *analytic continuation*. For example, Chiou and Youngs

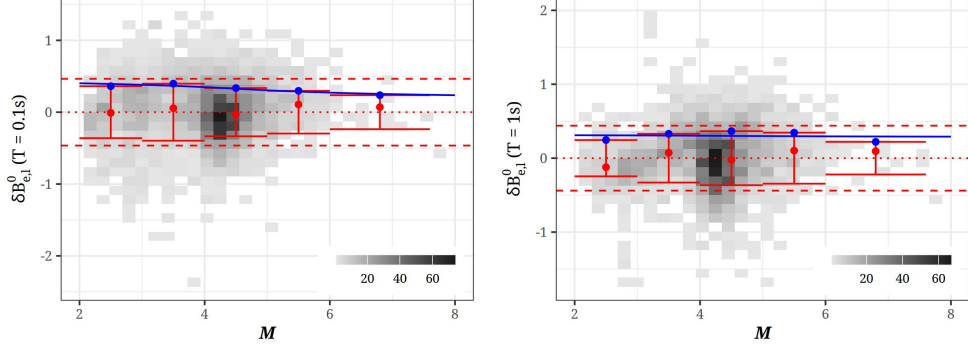


Fig. 7 Event-to-event random-effects (δB^0) versus M at $T = 0.1s$ (left-panel) and $T = 1s$ (right-panel). The grey scale presents the number of events in $\delta B^0_{e,l}$ M bins. The red markers and error-bars show bin-wise median and $\pm 0.7413 \cdot \text{IQR}$ of δB^0 , respectively. The dashed red lines delimit the $\pm \tau_0$. The blue dots show the bin-wise $0.7413 \cdot \text{IQR}$, while the blue line shows the heteroscedastic $\tau_0(M)$ fit to these values.

(2014) uses hyperbolic trigonometric functions in their GMPM fixed-effects for similar reasons.

$$\tau_0(M) = \tau_0(M_{ref}) + d_1 \cdot \text{arcsinh}(M - M_{ref}) \quad (6)$$

The blue lines in Figure 7 are the heteroscedastic $\tau_0(M)$ fitting to the bin-wise $0.7413 \cdot \text{IQR}$ values. While the $\tau_0(M)$ for $T = 0.1s$ follows the decreasing trend towards larger M , that for $T = 1s$ is rather *flat*. It is to be noted that the studentised BP test only tests the heterogeneity of variance with respect to independent variables, but does not imply monotonic increase/decrease with it.

5.2.5 Site variability

Site variability ($\Delta S2S$) random-effects capture the empirical site-to-site response variability with respect to a hypothetical reference condition. Generally, the reference site conditions in Eurocode8 are considered to be those with $V_{s30} = 800m.s^{-1}$. However, there are very few sites in ground-motion datasets with either a reliable, measured V_{s30} or the response of sites with $V_{s30} = 800m.s^{-1}$ is highly variable. The datasets used in derivation of this GMPM has 1072 stations with measured V_{s30} out of 2988 sites for which $\delta S2S_s$ could be estimated in the mixed-effects regression. Including V_{s30} in the GMPM's fixed-effects formula would force the regression to consider only data with associated V_{s30} . Therefore, as in Kotha et al. (2016) and K20, we validate the physical meaning of $\Delta S2S$ by plotting them again their available measured V_{s30} in Figure 8.

The description of Figure 8 is the same as that of 7. Unlike with $\Delta B^0 - M$, in this case, the bin-wise median of $\Delta S2S$ is away from zero baseline. The bins are chosen to imitate the Eurocode8 classification (EC8 class): class A for $V_{s30} \geq 800m.s^{-1}$,

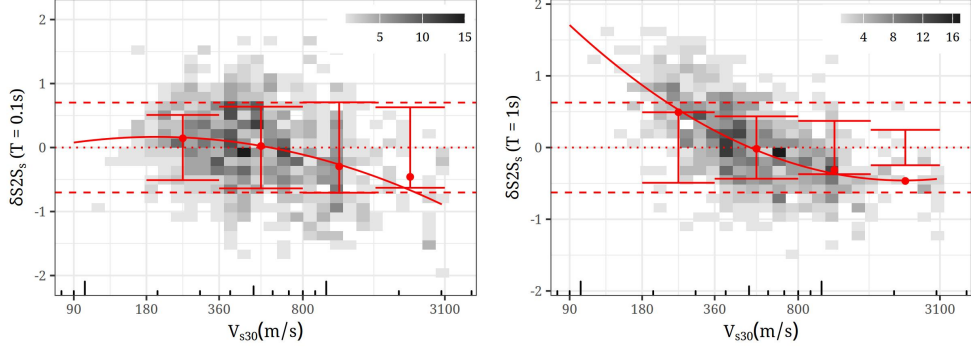


Fig. 8 Site-to-site random-effects ($\Delta S2S$) versus M at $T = 0.1s$ (left-panel) and $T = 1s$ (right-panel). The grey scale presents the number of sites in $\delta S2S_s$ V_{s30} bins. The red markers and error-bars show bin-wise median and $\pm 0.7413 \cdot IQR$ of $\Delta S2S$, respectively. The dashed red lines delimit the $\pm \phi_{S2S}$. The red line follows the quadratic site-response function $SR(V_{s30})$

class B for $360 \leq V_{s30} < 800 m.s^{-1}$, and class C for $180 \leq V_{s30} < 360 m.s^{-1}$. Class D for $V_{s30} < 180 m.s^{-1}$ is not shown because of the very sparse sampling of such of sites in the dataset. An additional bin with $V_{s30} \geq 1600 m.s^{-1}$ is added to the plot to distinguish *very hard rock* sites from class A. The red markers and the red error-bars showing the bin-wise (or EC8 class) medians and $0.7413 \cdot IQR$ of $\Delta S2S$, respectively, show very different trends for $T = 0.1s$ and $T = 1s$. At short-periods, $T = 0.1s$ in the left-panel, there appears to be a very weak scaling of (site amplification) $\Delta S2S$ with V_{s30} , with the bin-wise medians rather close to zero. The $0.7413 \cdot IQR$ values, however, are largest for $V_{s30} \geq 800 m.s^{-1}$, and lowest for $180 \leq V_{s30} < 360 m.s^{-1}$ sites. The bin-wise medians and $IQRs$ suggest that the short-period site-responses cannot be explained by V_{s30} , and that the response of *reference rock* sites with $V_{s30} \geq 800 m.s^{-1}$ is too large to be considered as 'reference' site condition. At long-periods, e.g. $T = 1s$ in the right-panel, the median and IQR trends reverse: with strong amplification and highest site-response variability towards $180 \leq V_{s30} < 360 m.s^{-1}$. Although one can argue that V_{s30} is an insufficient parameter to describe site-response adequately, the trends seen in Figure 8 support the claim that $\Delta S2S$, despite being statistical (random-effects) estimates, have a physical meaning (Kotha et al. 2018).

$$SR(V_{s30}) = g_0 + g_1 \cdot \ln(V_{s30}/800) + g_2 \cdot (\ln(V_{s30}/800))^2 \quad (7)$$

This study estimates $\delta S2S_s$ and its epistemic uncertainty for about 3400 sites. Sites with sufficient number of records (e.g. more than 10) have epistemic uncertainty small enough that their $\delta S2S_s$ can be used in site-specific PSHA. For sites without enough data, the $\Delta S2S$ can be first correlated to site-response proxies, and the empirical relations can be used to make site-response predictions, e.g. clustered in to site-classes (Kotha et al. 2018), correlated to a site-response proxy (Loviknes et al. 2024; Weatherill et al. 2023), spatially smoothed using a geospatial tool (Parker and Baltay 2022; Chakraborty and Kotha 2026). In lieu of a fully data-driven uncertainty projected site-response map from this dataset (Chakraborty et al. 2026), we fit $\Delta S2S$ V_{s30} quadratic function (Equation (7)) similar to that in K20. Preliminary

1151 site-response maps are shown in Figures S3 in supplementary materials for reference.
1152

1153 The red curves in Figure 8 are Equation (7) fits, which coincide very well with
1154 the bin-wise medians (red dots). Figure 8 suggests that the site-response variability
1155 is also heteroscedastic with V_{s30} . But we decided not to overly complicate the model
1156 given that only a third of all sites in the dataset have a measured V_{s30} . The period-
1157 dependant coefficients g_0 , g_1 and g_2 from Equation (7) are shown in Figure S5 of the
1158 supplementary materials. Nevertheless, a new site-response model can be derived from
1159 $\Delta S2S$ depending on the application (Bindi et al. 2017; Weatherill et al. 2023).
1160

1161 5.2.6 Residual variability

1162 The *left-over* residuals capture the record-to-record variability filtered for all mixed-
1163 effects. These residuals $\epsilon = \mathcal{N}(0, \phi_0^2)$ are still considered only *apparently* aleatory.
1164 A few studies have shown that the ϵ carry signatures of secondary phenomena,
1165 such as the shear-wave radiation pattern, directivity, nonlinear soil response, etc.
1166 The fully non-ergodic, varying coefficient, cell-based attenuation models reviewed in
1167 Lavrentiadis et al. (2023) show a substantial reduction in ϕ_0 compared to simpler,
1168 mixed-effects parametric GMPMs such as the one in this study. This means that, ϵ
1169 contain more repeatable effects that can be modelled as mixed-effects in the next
1170 generation of GMPMs. The Kotha and Traversa (2024) study presented an $[\mathbf{M}, R, T]$
1171 dependence of ϕ_0 , but without a clear trend that can be analytically modelled and
1172 used in predictions. In this study, we have observed the $[\mathbf{M}, R, T]$ dependence of ϕ_0 ,
1173 and that the ϵ do not show any $\mathbf{M}$ - and R -dependencies, assuring that the fixed-
1174 effects are unbiased. Figure S4 in supplementary materials shows the variation of ϵ
1175 with distance.
1176

1178 6 Applicability of the model

1180 The applicability of the new GMPM is T dependent. Figure 9 shows the
1181 100th, 50th, 25th quartiles of $\mathbf{M}$ (left-panel) and R (right-panel) ranges of the dataset
1182 used in regression of $SA(T)$. Compared to the $K20$ the new model has a lower
1183 minimum (25th in left-panel) and a higher maximum (100th in left-panel) $\mathbf{M}$ of appli-
1184 cability, but only up to $T = 2s$. At $T > 2s$, the amount of data from $\mathbf{M} \leq 4$ drops
1185 rapidly due to the poor signal-to-noise ratio of records, mostly in BELSHAKE and
1186 EPOS-FR datasets. Therefore, we suggest this model be used only for $\mathbf{M} \geq 4$ for
1187 $SA(T > 2s)$ predictions. In the right-panel, the lower limit (25th) of applicability of
1188 the new model is around 50km, while that of $K20$ is below 30km. However, Figure
1189 4 shows that the epistemic uncertainty - σ_μ - of the new model is lower than that of
1190 $K20$ at all distances. Therefore, the only precaution is that the new GMPM may not
1191 be used reliably for $\mathbf{M} < 4$ for $SA(T > 2s)$ predictions.
1192

1193 $K20$ elaborates on the application of the GMPM for ergodic, regional, and site-
1194 specific predictions. The same would apply for this GMPM as well. Subsequently,
1195 Weatherill et al. (2020) adapted the $K20$ and its Kotha et al. (2022b) update as
1196 a regionalised backbone GMPM logic-tree for ESHM20 in an application oriented

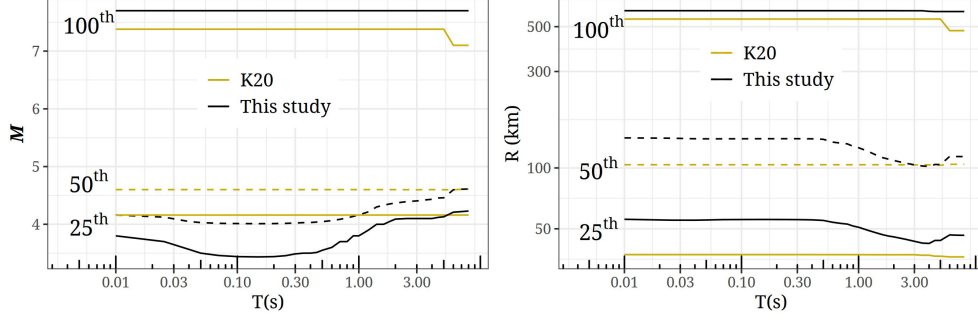


Fig. 9 The $[25^{th}, 50^{th}, 100^{th}]$ quantiles of M and R (km) in the datasets used to regress $K20$ (yellow) and the GMPM in this study (black)

manner. Similarly, the new GMPM is being adapted for the national seismic hazard map of Belgium through personal communication with Kris Vanneste (BELSHAKE project), and for the hazard map of France (Alceste project).

The GMPM is implemented as a Python script compatible with [OpenQuake](#), provided as an electronic supplement here, and will be integrated into the Ground-Shaking Intensity Model ([GSIM](#)) module of OpenQuake ([Pagani et al. 2014](#)). Three versions of the model will be made available in OpenQuake : 1) a **Base** version with no attenuation ($\delta c_3^s = 0$ and $\delta c_3^e = 0$), locality ($\delta L2L_l = 0$), and site-specific ($\delta S2S_s = 0$) or site-response adjustments, 2) a **Site** version with no attenuation and locality adjustments, but with the V_{s30} scaled site-response component (Equation (7)), and (3)) a **Regional** version with event (e) and site (s) coordinate dependent attenuation (Figure 5) and locality (Figure 6) adjustments, and with the V_{s30} scaled site-response component (Equation (7)). A fourth model upgrading the **Regional** version with uncertainty projected site-response map by [Chakraborty et al. \(2026\)](#) will be integrated eventually. The electronic supplement presents preliminary version of the envisioned site-response maps in Figures S3 for reference. Casting the model as backbone logic-tree is possible, but that will depend on the application. To perform a Bayesian update of the GMPM with new data, and to develop an application oriented version, we encourage users to contact us directly.

7 Discussion

The new GMPM supersedes the $K20$ GMPM for shallow crustal earthquakes in Euro-Mediterranean region. Although the new GMPM is not substantially more sophisticated than $K20$, it is worth evaluating refinements in comparison to the choices made in $K20$. For this purpose, we have performed several tests but have not reported them here for the sake of brevity. For example, as reported in [Laurendeau and Kotha \(2025\)](#), homogenization of M has substantially reduced the τ_0 . The improvements in M -scaling towards $M \leq 4.5$ in Figure 2, the reduction in $[M, R, T]$

dependent within-model epistemic uncertainty (σ_μ) is evident from Figure 4, the changes to aleatory variability models is shown in Figure 3 and Figure 7. We have not attempted to assign physical meaning to the attenuation and locality random-effects in Figure 5 and 6, respectively. Such an analyses is deemed more appropriate with a Fourier version of the GMPM in an ongoing study. The GMPM has wider $[M, R, T]$ applicability (Figure 9), can be updated in a Bayesian framework, and adaptable to ergodic, regional or site-specific hazard assessments. What remains to be discussed is the quantitative improvements from using a grid-based regionalisation, and the introduction of event location based attenuation (Figure 1).

7.1 Improvements with respect to K20

To quantitatively validate the improvements from refining the regionalisation models in the GMPM, we regressed three additional models, each incorporating one additional improvement starting from $K20$ as the baseline. The baseline GMPM is a Bayesian version of $K20$: the regressed dataset is identical to that used in $K20$ but with homogenized M - labelled as $D20$; the attenuation regions and tectonic localities identical to those of $K20$ - labelled as $R20$; and the mixed-effects identical to that of $K20$ - labelled as $M20$. The baseline GMPM - Bayesian version of $K20$ - is labelled as $D20 - R20 - M20$. The second model replaces the dataset with the compendium dataset from this study - labelled as $D26$. The second GMPM is thus labelled $D26 - R20 - M20$. The third GMPM, in addition, replaces the regionalisation models with grid-based regionalisation - $R26$ - but without changing the mixed-effects formula; i.e. the attenuation variability is only dependent on station location and not on event location. The third GMPM is labelled $D26 - R26 - M20$. The final GMPM is the one presented in this study: with the $D26$ dataset and the $R26$ grid-based regionalisation models, and the additional event location based attenuation variability. The final models mixed-effects formula contains one additional random effect - labelled as $M26$. Therefore, the GMPM presented in this study is labelled here as $D26 - R26 - M26$.

All four models are regressed for the entire $SA(T = 0.01 - 8s)$ range, and a comparison of their mixed-effects is shown in Figure 10. The description of Figure 10 is similar to Figure 3. The comparison revealed negligible differences in the fixed-effects of the three latter models: $D26 - R20 - M20$, $D26 - R26 - M20$, and $D26 - R26 - M26$. The $D20 - R20 - M20$ is the old one, much like in the comparison of $K20$ to the new GMPM in this study. When moving from $D20$ to $D26$, the quadratic M -scaling coefficient b_2 (Equation (1)) becomes negative, unlike the $D20 - R20 - M20$ model. This implies that the more *physically meaningful* and *theoretically consistent* non-positive b_2 of the three $D26$ based models is driven by the increased of events and sites included in the $D26$ dataset. We infer that a reliable calibration of M -scaling for small-moderate sized earthquakes - which are highly relevant to low-moderate seismicity regions - requires homogenized M of numerous $M \leq 4.5$. It is worth reiterating that, $M \leq 4.5$ are indispensable in development of more sophisticated and widely applicable fully non-ergodic GMPMs (Lavrentiadis

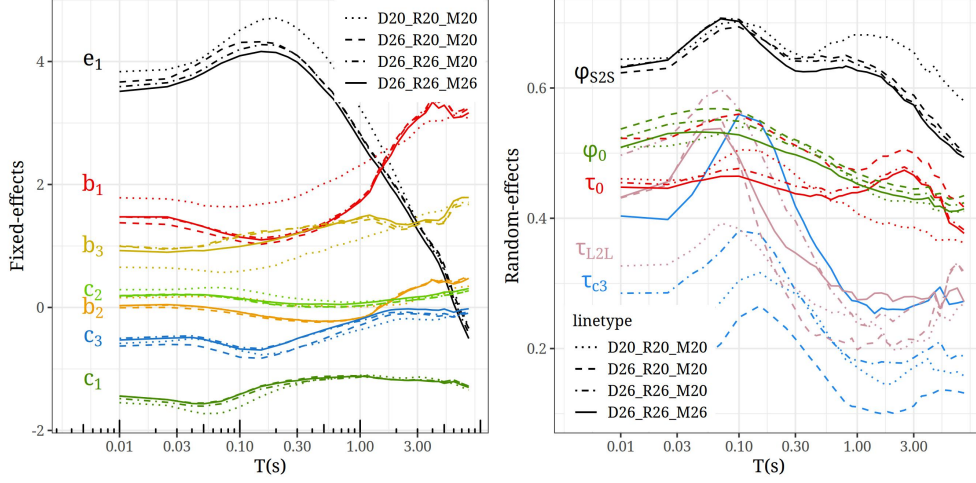


Fig. 10 Frequency dependence of fixed-effects coefficients and random-effects standard-deviations of four intermediate models for quantitative validation of improvements. *D20* represents the ESM18 flatfile against the combined flatfile from this study (*D26*). *R20* for the ESHM20 regionalisation and *R26* for grid cell regionalisation. *D20* for *K20* model and *D26* for the model of the study.

et al. 2023). The other M -scaling coefficient b_3 in Equation (1), describing the $SA(T)$ scaling for larger events with $M > 5.7$, is more positive (and equal) among the three latter GMPMs regressed over *D26* than that of *D20* – *R20* – *M20*; very likely due to additional $M > 5.7$ data and the correction to b_2 . Also worth noting is that the fixed-effects coefficients of Equation (3) are practically identical among all four models.

Table 1 provides a comparison of the *D20* and *D26* datasets, and *R20* and *R26* regionalisation models. Comparing the random-effects standard-deviations in Figure 10, the residual (ϕ_0) and between-site (ϕ_{S2S}) standard-deviations of the four models are very similar, despite the *D26* containing three times the number of records, and twice the number of sites in *D20*. While the record-to-record aleatory variability ϕ_0 is unaffected, the between-event τ_0 of *D26* – *R20* – *M20* is larger than the rest; followed by *D20* – *R20* – *M20*, *D26* – *R26* – *M20*, and *D26* – *R26* – *M26*. Consequently, the non-ergodic aleatory $\sigma_0 = \sqrt{\tau_0^2 + \phi_0^2}$ of the *D26* – *R26* – *M26* would be the lowest, and that of *D26* – *R20* – *M20* would be the highest among the four models. This means that, transition from *R20* to *R26* regionalisation models is necessary to mitigate the increase in σ_0 . This inference is also supported by the larger τ_{L2L} of *D26* – *R26* – *M20* and *D26* – *R26* – *M26* models, which is introduced to quantify repeatable event effects in tectonic localities, and be treated as an epistemic uncertainty in a ground-motion logic-trees. τ_{L2L} of *D20* – *R20* – *M20* is the lowest because it was regressed on a much smaller dataset, in terms of both diversity and resolution of the tectonic locality regionalisation models. The same reasoning can be used to explain the larger $\tau_{c3} = \sqrt{\tau_{c3,e}^2 + \tau_{c3,s}^2 + 2\text{Cov}(\Delta c_3^e, \Delta c_3^s)}$ of *D26* – *R26* – *M26* and $\tau_{c3} = \tau_{c3,s}$ of *D26* – *R26* – *M20* models. Among these two, τ_{c3} of *D26* – *R26* – *M26*

1335 is larger due to its additional event location dependent Δc_3^e , which distinguishes its
1336 mixed-effects formula as $M26$. Generally, a larger random-effects standard-deviation
1337 suggests that the random-effects grouping is more appropriate for the dataset (Bates
1338 et al. 2015). In a way, these observations support each of the refinements going from
1339 $D20 - R20 - M20$ to $D26 - R26 - M26$, which validates our proposal to supersede
1340 $K20$ with the new GMPM from this study. It is important to note that, the new
1341 GMPM - if cast as a backbone ground-motion logic-tree - will have greater epistemic
1342 uncertainties due to its larger τ_{L2L} and τ_{c3} ; but only for those attenuation regions
1343 and localities where no data is available. Therefore, it is strongly recommended that
1344 the Bayesian GMPM is updated wherever data becomes available.

1345

1346 Introducing a source related attenuation Δc_3^e is certainly effecting the estimation
1347 of the locality-to-locality spatial variability ($\Delta L2L$) which could explain the large
1348 difference between the τ_{L2L} between the $D26 - R26 - M20$ and $D26 - R26 - M26$
1349 models. The increase in τ_{c3} is compensated by the decrease in τ_{L2L} in the last two
1350 models.

1351

1352 7.2 Improvements in Information Criteria

1353

1354 $K20$ performed a k -fold cross-validation test on their *frequentist* robust linear
1355 mixed-effects GMPM to justify the introduction of ΔB^0 , $\Delta S2S$, Δc_3^e and $\Delta L2L$
1356 random-effects groups. For the four Bayesian GMPMs evaluated here, to avoid
1357 repeated refitting of models in a k -fold cross-validation, we employ the computation-
1358 ally efficient and fully *Bayesian* Watanabe-Akaike Information Criterion (*WAIC*).
1359 *WAIC* is an asymptotic approximation of k -fold cross-validation. It is a statistical
1360 measure designed to evaluate how well a model predicts data it was not trained on,
1361 particularly in Bayesian models. It estimates a model's predictive accuracy for new
1362 data, generalizing the Akaike Information Criterion (AIC) to complex models. Unlike
1363 traditional criteria like AIC or DIC (Deviance Information Criterion), *WAIC* uses
1364 the entire posterior distribution of model parameters, making it suitable for complex
1365 models. *WAIC* estimates the expected out-of-sample predictive accuracy by combin-
1366 ing the model's fit to the observed data with a penalty for model complexity. Lower
1367 *WAIC* values indicate better predictive performance.

1368

1369

1370

1371

1372

1373

1374

1375

1376

1377

1378

1379

1380

Table 2 Statistical criteria *WAIC* (Watanabe-Akaike Information Criterion) for comparing the three test cases with final dataset, at four different periods : 0.1, 0.3, 1 and 3s

	<i>WAIC</i>			
	T = 0.1s	0.3s	1s	3s
D26-R20-M20	99 163	90 344	72 183	58 967
D26-R26-M20	96 329	88 541	72 626	57 103
D26-R26-M26	92 517	86 802	70 989	53 853

The *WAIC* are presented in Table 2 for the three models with compendium dataset *D26* used in this study. The *WAIC* is not reliable for comparing models with different dataset, i.e. *D20*, but we have demonstrated the importance of improving dataset in earlier sections. For each modification in the model, i.e. *R26* and *M26*, the *WAIC* decreases for all *T*. This means that, each additional complexity in the model captures more information and does not over fit the data. We have also tested the log-marginal likelihood (*LLH*), which is the logarithm of the probability of the data in the model, taking into account all possible parameters weighted by the a priori probability. Used only to compare Bayesian models, it balances the best likelihood with the lowest model complexity. The *LLH* results align and confirm the improvements in *WAIC*.

The resulting compendium dataset with homogenised *M* is shared as an electronic supplement. While the GMPM predicts *SA(T)*, the dataset, with random-effects and residuals included, has been used to develop spatial cross-correlation models (Dami-ani et al. 2026a,b). These correlation models can be used in vector-valued PSHA, prediction of advanced intensity measures such as the Average Spectral Acceleration (*AvgSA*, Kohrangi et al. (2018)), producing ShakeMaps (Cauzzi et al. 2018), and performing seismic risk assessment of distributed assets and portfolios (Park et al. 2007). The fully consistent uncertainty projected site-response maps developed from the site-specific terms of the GMPM, and the mixed-effects of the GMPM can be updated as and when new ground-motion data becomes available.

8 Perspectives and Conclusions

One of the main areas for improvement is the dataset. The model covers larger areas and is more accurate, particularly in regions with limited data. However, it remains a critical point of the regression. Estimating the distance and, in particular, the depth of the earthquake can be improved. This is because the model uses only three discrete depth zones. *M* estimation is another limitation of the dataset due to differences in their computation between datasets, as well as limitations in calculating *M* for small-moderate sized earthquake recordings contaminated by low-frequency anthropogenic and natural noise. As *M* estimation requires a more accurate crustal models, this model could overcome this limitation. The EPOS-ESM dataset is extensive for Europe, but as it does not cover some areas, we had to add other datasets to cover a larger area. Combining the three datasets limits their reliability because the signal processing methods and event metadata in the datasets are often not homogeneous. We mitigated these issues by removing common records and using a homogenised magnitude, but creating a higher quality dataset would be an entirely different project. The distance metric can also be a limitation of the model. Indeed, we are using Joyner-Boore distance but the reality of the path can be different from the modeled distance used. Finally, even with the combination of three datasets, some regions still lack data. This model is not usable across Europe without a proxy, therefore, the immediate objective is to determine a widely available *proxy* for the

1427 random-effects of the GMPM, as explored in [Kotha et al. \(2022a\)](#).

1428

1429 Data distribution in Europe is highly heterogeneous. The division of Europe into
1430 cells reduces the problem of extrapolating data over large regions, but creates other
1431 issues. The uncertainty between cells can also be highly variable. For example, Italy
1432 has a large amount of data, so the number of data points is high. However, in western
1433 France, the number of data points in each cell can be very low.

1434

1435 We should also note that the attenuation kernel descriptions in this paper are
1436 greatly simplified compared to the paper of [Mayor et al. \(2014\)](#). As the ground-
1437 motions used to compute the response spectra comprises many different types of
1438 waves, we had to simplify and summarise the reality and complexity of wave sen-
1439 sitivity. This model now presents two random parameters to describe attenuation
1440 variability. We tested many alternative solutions for modelling attenuation, and the
1441 one presented here was the most reliable. While it may be slightly more complex
1442 to analyse the two attenuation variability maps together, they need to be treated
1443 as complementary. The difference found between the two maps could be another
1444 limitation. Although we have substantially improved the representation of attenua-
1445 tion physics, it still does not represent the propagation physics perfectly. Therefore,
1446 another critical improvement would be to refine the parametric form of the GMPM.

1447

1448 We present here a new Bayesian parameter GMPM to predict response spectra
1449 of shallow crustal earthquakes in Europe. Relying on the parametric form of [Kotha
1450 et al. \(2020\)](#), this model offers several significant improvements over the [Kotha et al.
1451 \(2020\)](#). The dataset covers a larger area, and especially presenting a much wider M
1452 range of applicability, which allows the model to be used in low-to-moderate seismic
1453 regions such as France and Belgium. Furthermore, the larger quantity of data enables
1454 larger regional event, site, attenuation, and tectonic variabilities to be estimated.
1455 Next improvement is the shift from expert-based regionalisation for Europe to a data-
1456 driven regionalisation. This has two main advantages: it allows attenuation variability
1457 in Europe to be analysed more accurately, and updated in a Bayesian framework.
1458 The third improvement is the combination of a physics-based and a GMPM to
1459 implement regionalisation of attenuation. Some random-effects residuals have been
1460 reduced, thereby improving the model. As we know, such models are important in
1461 seismic hazard studies, but we should also recognise the importance of accurate and
1462 local attenuation variability in estimating earthquakes magnitudes in dataset. This
1463 work is important for continuing to improve the regionalisation of the attenuation.
1464 First, this model should be derived in the Fourier amplitude domain. This will enable
1465 the results of attenuation variability to be compared with other geophysical data on
1466 the Earth. The main objective is to improve our understanding of attenuation, and
1467 eventually, to find a proxy that will help us to reduce the uncertainty in the model
1468 and improve the seismic hazard prediction.

1469

1470

1471

1472

Supplementary information. The OpenQuake implementation in python code and the regionalisation maps can be found https://github.com/PaulineGgs/G26-GMM-in-OpenQuake . The coefficient table, the regionalisation maps and the dataset are provided in supplementary materials.	1473 1474 1475 1476
Acknowledgements. This project was funded by the <i>Agence Nationale de la Recherche</i> (ANR - NGS-EU / ANR-22-CE56-0003-01). We also thank Kris Vanneste and the Alceste project for their feedback.	1477 1478 1479 1480
References	1481 1482 1483
Al Atik L (2015) NGA-East: Ground-Motion Standard Deviation Models for Central and Eastern North America. PEER Report No. 2015/07, Pacific Earthquake Engineering Research Center, University of California, Berkeley	1484 1485 1486
Atik LA, Youngs RR (2014) Epistemic Uncertainty for NGA-West2 Models. Earthquake Spectra 30(3):1301–1318. https://doi.org/10.1193/062813EQS173M	1487 1488 1489
Basili R, Brizuela B, Herrero A, et al (2021) The Making of the NEAM Tsunami Hazard Model 2018 (NEAMTHM18). Frontiers in Earth Science 8:616594. https://doi.org/10.3389/feart.2020.616594	1490 1491 1492 1493
Bates D, Mächler M, Bolker B, et al (2015) Fitting Linear Mixed-Effects Models Using lme4 . Journal of Statistical Software 67(1). https://doi.org/10.18637/jss.v067.i01	1494 1495 1496
Bergamo P, Sunny J, Grigoratos I, et al (2025) Modelling soil response at the national scale for Switzerland in the framework of risk assessment of induced seismicity. URL https://meetingorganizer.copernicus.org/EGU25/presentation/EGU25-8121 , EGU General Assembly 2025, EGU25-8121	1497 1498 1499 1500
Bindi D, Kotha SR (2020) Spectral decomposition of the Engineering Strong Motion (ESM) flat file: regional attenuation, source scaling and Arias stress drop. Bulletin of Earthquake Engineering 18(6):2581–2606. https://doi.org/10.1007/s10518-020-00796-1	1501 1502 1503 1504 1505
Bindi D, Cotton F, Kotha SR, et al (2017) Application-driven ground motion prediction equation for seismic hazard assessments in non-cratonic moderate-seismicity areas. Journal of Seismology 21:1201–1218. https://doi.org/10.1007/s10950-017-9661-5 , publisher: Springer ADS Bibcode: 2017JSeis..21.1201B	1506 1507 1508 1509 1510
Bindi D, Kotha SR, Weatherill G, et al (2019) The pan-European engineering strong motion (ESM) flatfile: consistency check via residual analysis. Bulletin of Earthquake Engineering 17(2):583–602. https://doi.org/10.1007/s10518-018-0466-x	1511 1512 1513 1514
Boore DM (2010) Orientation-Independent, Nongeometric-Mean Measures of Seismic Intensity from Two Horizontal Components of Motion. Bulletin of the Seismological Society of America 100(4):1830–1835. https://doi.org/10.1785/0120090400	1515 1516 1517 1518

1519 Breusch TS, Pagan AR (1979) A Simple Test for Heteroscedasticity and Random Coef-
1520 ficient Variation. *Econometrica* 47(5):1287–1294. <https://doi.org/10.2307/1911963>,
1521 publisher: [Wiley, Econometric Society]
1522

1523 Brune JN (1970) Tectonic stress and the spectra of seismic shear waves from earth-
1524 quakes. *Journal of Geophysical Research* 75(26):4997–5009. [https://doi.org/10.](https://doi.org/10.1029/JB075i026p04997)
1525 [1029/JB075i026p04997](https://doi.org/10.1029/JB075i026p04997)
1526

1527 Buscetti M, Traversa P, Perron V, et al (2024) Epos-France Database of Earthquake
1528 Ground Motion for Mainland France: 1996–2021 Updated Version. *Seismological*
1529 *Research Letters* 96(2A):1214–1226. <https://doi.org/10.1785/0220240053>
1530

1531 Campillo M, Plantet JL (1991) Frequency dependence and spatial distribution of
1532 seismic attenuation in France: experimental results and possible interpretations.
1533 *Physics of the Earth and Planetary Interiors* 67(1):48–64. [https://doi.org/10.1016/](https://doi.org/10.1016/0031-9201(91)90059-Q)
1534 [0031-9201\(91\)90059-Q](https://doi.org/10.1016/0031-9201(91)90059-Q)
1535

1536 Cauzzi C, Clinton J, Faenza L, et al (2018) ShakeMapEU: Towards an Integrated Euro-
1537 pean ShakeMap System. URL [https://ui.adsabs.harvard.edu/abs/2018EGUGA..20.](https://ui.adsabs.harvard.edu/abs/2018EGUGA..20.9016C)
1538 [9016C](https://ui.adsabs.harvard.edu/abs/2018EGUGA..20.9016C), EGU General Assembly 2018, p. 9016
1539

1539 Chakraborty A, Goto H (2018) A Bayesian model reflecting uncertainties on map
1540 resolutions with application to the study of site response variation. *Geophysical*
1541 *Journal International* 214(3):2264–2276. <https://doi.org/10.1093/gji/ggy260>
1542

1543 Chakraborty A, Kotha SR (2026) Regional Site Response Model via Uncertainty
1544 Projected Mapping: Application to Kanto Region, Japan. *Earthquake Spectra*
1545 42(1):e70025. <https://doi.org/10.1002/esp4.70025>
1546

1547 Chakraborty A, P. G, Kotha SR (2026) Probabilistic, empirical, and scalable site
1548 response model for europe via uncertainty projected mapping, submitted to *Bulletin*
1549 *of Earthquake Engineering*
1550

1551 Chiou BSJ, Youngs RR (2014) Update of the Chiou and Youngs NGA Model for
1552 the Average Horizontal Component of Peak Ground Motion and Response Spectra.
1553 *Earthquake Spectra* 30(3):1117–1153. <https://doi.org/10.1193/072813EQS219M>
1554

1555 Cochran ES, Baltay A, Chu S, et al (2024) SCEC/USGS Community Stress-Drop
1556 Validation Study: How Spectral Fitting Approaches Influence Measured Source
1557 Parameters. *Bulletin of the Seismological Society of America* 115(3):760–776. <https://doi.org/10.1785/0120240140>
1558

1559 Cornou C, Ampuero J, Aubert C, et al (2021) Rapid response to the M_w 4.9 earthquake of November 11, 2019 in Le Teil, Lower Rhône Valley, France.
1560 *Comptes Rendus Géoscience* 353(S1):441–463. <https://doi.org/10.5802/crgeos.30>
1561
1562
1563
1564

Cotton F, Scherbaum F, Bommer JJ, et al (2006) Criteria for Selecting and Adjusting Ground-Motion Models for Specific Target Regions: Application to Central Europe and Rock Sites. <i>Journal of Seismology</i> 10(2):137–156. https://doi.org/10.1007/s10950-005-9006-7	1565 1566 1567 1568 1569
Damiani A, Kohrangi M, Kotha SR, et al (2026a) Ergodic and quasi-non-ergodic spatial cross-period correlation models for spectral accelerations of pan-European seismicity, submitted to <i>Earthquake Engineering and Structural Dynamics</i>	1570 1571 1572 1573
Damiani A, Kohrangi M, Kotha SR, et al (2026b) A hybrid approach for spatial cross-IM correlation modelling merging principal component analysis and geostatistical coregionalization, submitted to <i>Earthquake Engineering and Structural Dynamics</i>	1574 1575 1576 1577
Danciu L, Nandan S, Reyes C, et al (2021) The 2020 update of the European Seismic Hazard Model - ESHM20: Model Overview. Tech. rep., EFEHR, https://doi.org/10.12686/A15 , eFEHR Technical Report 001, v1.0.0	1578 1579 1580 1581
Danciu L, Giardini D, Weatherill G, et al (2024) The 2020 European seismic hazard model: overview and results. <i>Natural Hazards and Earth System Sciences</i> 24(9):3049–3073. https://doi.org/https://doi.org/10.5194/nhess-24-3049-2024 , publisher: Copernicus GmbH	1582 1583 1584 1585
Dawood HM, Rodriguez-Marek A (2013) A method for including path effects in ground-motion prediction equations: An example using the mw 9.0 tohoku earthquake aftershocks. <i>Bulletin of the Seismological Society of America</i> 103(2B):1360–1372. https://doi.org/10.1785/0120120125	1586 1587 1588 1589 1590
Douglas J (2024) Ground motion prediction equations 1964-2023 (incomplete). URL https://scholar.google.com/citations?view_op=view_citation&hl=en&user=Pk9ARLIAAAAJ&sortby=pubdate&citation_for_view=Pk9ARLIAAAAJ:WIXB4To3Tx4C , figshare	1591 1592 1593 1594 1595
Douglas J, Aochi H (2016) Assessing components of ground-motion variability from simulations for the Marmara Sea region (Turkey). <i>Bulletin of the Seismological Society of America</i> 106(1):300–306. https://doi.org/https://doi.org/10.1785/0120150177 , publisher: Seismological Society of America	1596 1597 1598 1599 1600
Faccioli E, Paolucci R, Vanini M (2015) Evaluation of Probabilistic Site-Specific Seismic-Hazard Methods and Associated Uncertainties, with Applications in the Po Plain, Northern Italy. <i>Bulletin of the Seismological Society of America</i> 105(5):2787–2807. https://doi.org/10.1785/0120150051	1601 1602 1603 1604 1605
Fukushima Y (1996) Scaling relations for strong ground motion prediction models with M 2 terms. <i>Bulletin of the Seismological Society of America</i> 86(2):329–336. https://doi.org/https://doi.org/10.1785/BSSA0860020329	1606 1607 1608 1609 1610

1611 Grad M, Tiira T, ESC Working Group (2009) The Moho depth map of the European
 1612 Plate. *Geophysical Journal International* 176(1):279–292. [https://doi.org/10.1111/](https://doi.org/10.1111/j.1365-246X.2008.03919.x)
 1613 [j.1365-246X.2008.03919.x](https://doi.org/10.1111/j.1365-246X.2008.03919.x)
 1614
 1615 Gregor N, Abrahamson NA, Atkinson GM, et al (2014) Comparison of NGA-
 1616 West2 GMPEs. *Earthquake Spectra* 30(3):1179–1197. [https://doi.org/10.1193/](https://doi.org/10.1193/070113EQS186M)
 1617 [070113EQS186M](https://doi.org/10.1193/070113EQS186M)
 1618
 1619 Grünthal G, Wahlström R (2012) The European-Mediterranean Earthquake Catalogue
 1620 (EMEC) for the last millennium. *Journal of Seismology* 16(3):535–570. [https://doi.](https://doi.org/10.1007/s10950-012-9302-y)
 1621 [org/10.1007/s10950-012-9302-y](https://doi.org/10.1007/s10950-012-9302-y)
 1622
 1623 Gómez-Rubio V (2020) Bayesian inference with INLA. Chapman & Hall/CRC Press,
 1624 Boca Raton, FL, URL <http://becarioprecario.bitbucket.io/inla-gitbook/index.html>
 1625
 1626 Hanks TC, Kanamori H (1979) A moment magnitude scale. *Journal of*
 1627 *Geophysical Research: Solid Earth* 84(B5):2348–2350. [https://doi.org/10.1029/](https://doi.org/10.1029/JB084iB05p02348)
 1628 [JB084iB05p02348](https://doi.org/10.1029/JB084iB05p02348)
 1629
 1630 Heller G, Sèbe O, Margerin L, et al (2025) Generalized inversion of source, site and
 1631 attenuation parameters using the radiative transfer theory: application to a French
 1632 dataset. <https://doi.org/10.5194/egusphere-egu25-10969>, EGU General Assembly
 1633 2025, EGU25-10969
 1634
 1635 Kale Ö, Akkar S, Ansari A, et al (2015) A Ground-Motion Predictive Model for
 1636 Iran and Turkey for Horizontal PGA, PGV, and 5% Damped Response Spectrum:
 1637 Investigation of Possible Regional Effects. *Bulletin of the Seismological Society of*
 1638 *America* 105(2A):963–980. <https://doi.org/10.1785/0120140134>
 1639
 1640 Kohrangi M, Kotha SR, Bazzurro P (2018) Ground-motion models for average spectral
 1641 acceleration in a period range: direct and indirect methods. *Bulletin of Earthquake*
 1642 *Engineering* 16(1):45–65. <https://doi.org/10.1007/s10518-017-0216-5>
 1643
 1644 Koller M (2016) **robustlmm** : An *R* Package for Robust Estimation of Linear
 1645 Mixed-Effects Models. *Journal of Statistical Software* 75(6):1–24. [https://doi.org/](https://doi.org/10.18637/jss.v075.i06)
 1646 [10.18637/jss.v075.i06](https://doi.org/10.18637/jss.v075.i06)
 1647
 1648 Kotha SR, Traversa P (2024) A Bayesian update of Kotha et al. (2020) ground-motion
 1649 model using Résif dataset. *Bulletin of Earthquake Engineering* 22(4):2267–2293.
 1650 <https://doi.org/10.1007/s10518-023-01853-1>
 1651
 1652 Kotha SR, Bindi D, Cotton F (2016) Partially non-ergodic region specific GMPE
 1653 for Europe and Middle-East. *Bulletin of Earthquake Engineering* 14(4):1245–1263.
 1654 <https://doi.org/10.1007/s10518-016-9875-x>
 1655
 1656 Kotha SR, Bindi D, Cotton F (2017) From Ergodic to Region- and Site-Specific
 Probabilistic Seismic Hazard Assessment: Method Development and Application at

European and Middle Eastern Sites. Earthquake Spectra 33(4):1433–1453. <https://doi.org/10.1193/081016eqs130m> 1657
1658
1659

Kotha SR, Cotton F, Bindi D (2018) A new approach to site classification: Mixed-effects Ground Motion Prediction Equation with spectral clustering of site amplification functions. Soil Dynamics and Earthquake Engineering 110:318–329. <https://doi.org/https://doi.org/10.1016/j.soildyn.2018.01.051>, publisher: Elsevier 1660
1661
1662
1663
1664

Kotha SR, Weatherill G, Bindi D, et al (2020) A regionally-adaptable ground-motion model for shallow crustal earthquakes in Europe. Bulletin of Earthquake Engineering 18(9):4091–4125. <https://doi.org/10.1007/s10518-020-00869-1> 1665
1666
1667
1668

Kotha SR, Bindi D, Cotton F (2022a) A regionally adaptable ground-motion model for fourier amplitude spectra of shallow crustal earthquakes in Europe. Bulletin of Earthquake Engineering 20(2):711–740. <https://doi.org/10.1007/s10518-021-01255-1> 1669
1670
1671
1672

Kotha SR, Weatherill G, Bindi D, et al (2022b) Near-source magnitude scaling of spectral accelerations: analysis and update of kotha et al. (2020) model. Bulletin of Earthquake Engineering 20(3):1343–1370. <https://doi.org/10.1007/s10518-021-01308-5> 1673
1674
1675
1676
1677

Kowsari M, Halldorsson B, Hrafnkelsson B, et al (2019) Calibration of ground motion models to Icelandic peak ground acceleration data using Bayesian Markov Chain Monte Carlo simulation. Bulletin of Earthquake Engineering 17(6):2841–2870. <https://doi.org/10.1007/s10518-019-00569-5> 1678
1679
1680
1681
1682

Kuehn N (2021) A Primer for using INLA to Estimate Ground-Motion Models. <https://doi.org/10.31224/osf.io/6ut3p> 1683
1684
1685

Kuehn NM, Abrahamson NA (2017) The Effect of Uncertainty in Predictor Variables on the Estimation of Ground-Motion Prediction Equations. Bulletin of the Seismological Society of America 108(1):358–370. <https://doi.org/10.1785/0120170166> 1686
1687
1688
1689

Kuehn NM, Kishida T, AlHamaydeh M, et al (2020) A Bayesian model for truncated regression for the estimation of empirical ground-motion models. Bulletin of Earthquake Engineering 18(14):6149. <https://doi.org/10.1007/s10518-020-00943-8> 1690
1691
1692
1693

Lammers S, Weatherill G, Grünthal G, et al (2019) EMEC-2014 - The European-Mediterranean Earthquake Catalogue – Version 2014. <https://doi.org/10.5880/GFZ.EMEC.2014.001>, GFZ Data Services 1694
1695
1696
1697

Lanzano G, Sgobba S, Luzi L, et al (2019) The pan-European Engineering Strong Motion (ESM) flatfile: compilation criteria and data statistics. Bulletin of Earthquake Engineering 61(3):343–356. <https://doi.org/10.1007/s10518-018-0480-z> 1698
1699
1700
1701
1702

1703 Lanzano G, Felicetta C, Pacor F, et al (2020) Methodology to identify the reference
 1704 rock sites in regions of medium-to-high seismicity: an application in Central Italy.
 1705 Geophysical Journal International 222(3):2053–2067. [https://doi.org/10.1093/gji/](https://doi.org/10.1093/gji/ggaa261)
 1706 [ggaa261](https://doi.org/10.1093/gji/ggaa261)
 1707
 1708 Laurendeau A, Kotha SR (2023) Moment-Magnitude Definition for Pan-European
 1709 Shallow Crustal Earthquakes: Impact on Fourier Ground-Motion Variability. URL
 1710 <https://asn.hal.science/irs-n-04171343>, paper presented at the 28th IUGG General
 1711 Assembly, Berlin, Germany, July 2023
 1712
 1713 Laurendeau A, Kotha SR (2025) Impact of mw definition approach on fourier ground-
 1714 motion variability of shallow crustal earthquakes in europe. Journal of Seismology
 1715 29:127–145. <https://doi.org/10.1007/s10950-024-10254-z>, publisher: Springer Ver-
 1716 lag
 1717
 1718 Lavrentiadis G, Abrahamson NA, Nicolas KM, et al (2023) Overview and introduction
 1719 to development of non-ergodic earthquake ground-motion models. Bulletin of Earth-
 1720 quake Engineering 21(11):5121–5150. <https://doi.org/10.1007/s10518-022-01485-x>
 1721
 1722 Lindgren F, Rue H (2015) Bayesian Spatial Modelling with R-INLA. Journal of
 1723 Statistical Software 63:1–25. <https://doi.org/10.18637/jss.v063.i19>
 1724
 1725 Liou I, Abrahamson N (2024) Framework for aleatory variability and epistemic
 1726 uncertainty for the ground-motion characterization based on the level of simpli-
 1727 fication. Bulletin of the Seismological Society of America 115(1):296–314. <https://doi.org/10.1785/0120240141>
 1728
 1729 Loviknes K, Cotton F, Weatherill G (2024) Exploring inferred geomorphological
 1730 sediment thickness as a new site proxy to predict ground-shaking amplification
 1731 at regional scale: application to Europe and eastern Türkiye. Natural Hazards
 1732 and Earth System Sciences (NHESS) 24(4):1223–1247. [https://doi.org/10.5194/](https://doi.org/10.5194/nhess-24-1223-2024)
 1733 [nhess-24-1223-2024](https://doi.org/10.5194/nhess-24-1223-2024)
 1734
 1735 Manchuel K, Traversa P, Baumont D, et al (2018) The French seismic CATalogue
 1736 (FCAT-17). Bulletin of Earthquake Engineering 16(6):2227–2251. [https://doi.org/](https://doi.org/10.1007/s10518-017-0236-1)
 1737 [10.1007/s10518-017-0236-1](https://doi.org/10.1007/s10518-017-0236-1)
 1738
 1739 Margerin L, Planès T, Mayor J, et al (2016) Sensitivity kernels for coda-wave
 1740 interferometry and scattering tomography: theory and numerical evaluation in two-
 1741 dimensional anisotropically scattering media. Geophysical Journal International
 1742 204(1):650–666. <https://doi.org/10.1093/gji/ggv470>
 1743
 1744 Mascandola C, Sgobba S (2025) Ground motion pulse-like detection and classification:
 1745 combining different approaches for comprehensive metadata assignment supporting
 1746 ML techniques for engineering applications. URL [https://ui.adsabs.harvard.edu/](https://ui.adsabs.harvard.edu/abs/2025EGUGA..27.4379M/abstract)
 1747 [abs/2025EGUGA..27.4379M/abstract](https://ui.adsabs.harvard.edu/abs/2025EGUGA..27.4379M/abstract), EGU General Assembly 2025, EGU25-4379
 1748

Mascandola C, D'Amico M, Russo E, et al (2023) Toward a renewed data processing of the Engineering Strong Motion (ESM) database. https://doi.org/10.5194/egusphere-egu23-2210 , EGU General Assembly 2023, Vienna, Austria, 24–28 Apr 2023, EGU23-2210	1749 1750 1751 1752
Massey Jr. FJ (1951) The Kolmogorov-Smirnov Test for Goodness of Fit. <i>Journal of the American Statistical Association</i> 46(253):68–78. https://doi.org/10.1080/01621459.1951.10500769	1753 1754 1755 1756
Mayor J, Margerin L, Calvet M (2014) Sensitivity of coda waves to spatial variations of absorption and scattering: radiative transfer theory and 2-D examples. <i>Geophysical Journal International</i> 197(2):1117–1137. https://doi.org/10.1093/gji/ggu046	1757 1758 1759 1760
Mayor J, Traversa P, Calvet M, et al (2018) Tomography of crustal seismic attenuation in Metropolitan France: implications for seismicity analysis. <i>Bulletin of Earthquake Engineering</i> 16(6):2195–2210. https://doi.org/10.1007/s10518-017-0124-8	1761 1762 1763 1764
Pagani M, Monelli D, Weatherill G, et al (2014) OpenQuake Engine: An Open Hazard (and Risk) Software for the Global Earthquake Model. <i>Seismological Research Letters</i> 85:692–702. https://doi.org/10.1785/0220130087 , aDS Bibcode: 2014SeiRL..85..692P	1765 1766 1767 1768
Park J, Bazzurro P, Baker JW (2007) Modeling spatial correlation of ground motion intensity measures for regional seismic hazard and portfolio loss estimation. In: Kanda J, Takada T, Furuta H (eds) <i>Applications of Statistics and Probability in Civil Engineering</i> . Taylor & Francis, London, pp 1–8	1769 1770 1771 1772 1773
Parker GA, Baltay AS (2022) Empirical Map-Based Nonergodic Models of Site Response in the Greater Los Angeles Area. <i>Bulletin of the Seismological Society of America</i> 112(3):1607–1629. https://doi.org/10.1785/0120210175	1774 1775 1776 1777
Pasyanos M (2015) Validation of Attenuation Models for Ground Motion Applications in Central and Eastern North America. <i>Earthquake Spectra</i> 31(4):2281–2300. https://doi.org/10.1193/052714EQS074M	1778 1779 1780 1781
Rodriguez-Marek A, Cotton F, Abrahamson NA, et al (2013) A model for single-station standard deviation using data from various tectonic regions. <i>Bulletin of the seismological society of America</i> 103(6):3149–3163. https://doi.org/https://doi.org/10.1785/0120130030	1782 1783 1784 1785
Sandıkaya MA, Kale Ö, Akkar S, et al (2025) A simulation-based regional ground-motion model for Eastern Türkiye. <i>Bulletin of Earthquake Engineering</i> 23(3):1001–1015. https://doi.org/10.1007/s10518-024-02058-w	1786 1787 1788 1789
Satriano C (2022) SourceSpec – Earthquake source parameters from P- or S-wave displacement spectra. https://doi.org/10.5281/zenodo.6954238 , zenodo	1790 1791 1792 1793
	1794

1795 Stafford PJ (2019) Continuous integration of data into ground-motion models using
1796 Bayesian updating. *Journal of Seismology* 23(1):39–57. <https://doi.org/10.1007/s10950-018-9792-3>
1797
1798
1799 Thompson EM, Hearne M, Aagaard BT, et al (2025) Automated, near real-time
1800 ground-motion processing at the U.S. Geological Survey. *Seismological Research*
1801 *Letters* 96(1):538–553. <https://doi.org/10.1785/0220240021>
1802
1803 Traversa P, Maufroy E, Hollender F, et al (2020) RESIF RAP and RLBP Dataset
1804 of Earthquake Ground Motion in Mainland France. *Seismological Research Letters*
1805 91(4):2409–2424. <https://doi.org/10.1785/0220190367>
1806
1807 Trugman DT, Shearer PM (2018) Strong Correlation between Stress Drop and Peak
1808 Ground Acceleration for Recent M 1–4 Earthquakes in the San Francisco Bay Area.
1809 *Bulletin of the Seismological Society of America* 108(2):929–945. <https://doi.org/10.1785/0120170245>
1810
1811 Vanneste K, Onvani M (2025) Belgian ground-motion database (BELSHAKE) flatfile.
1812 Royal Observatory of Belgium, doi = <https://zenodo.org/records/15241694>
1813
1814 Walpersdorf A, Sue C, Baize S, et al (2015) Coherence between geodetic and seismic
1815 deformation in a context of slow tectonic activity (SW Alps, France). *Journal of*
1816 *Geodynamics* 85:58–65. <https://doi.org/10.1016/j.jog.2015.02.001>
1817
1818 Weatherill G, Kotha SR, Cotton F (2020) A regionally-adaptable “scaled backbone”
1819 ground motion logic tree for shallow seismicity in Europe: application to the 2020
1820 European seismic hazard model. *Bulletin of Earthquake Engineering* 18(11):5087–
1821 5117. <https://doi.org/10.1007/s10518-020-00899-9>
1822
1823 Weatherill G, Crowley H, Roullé A, et al (2023) Modelling site response at regional
1824 scale for the 2020 European Seismic Risk Model (ESRM20). *Bulletin of Earthquake*
1825 *Engineering* 21(2):665–714. <https://doi.org/10.1007/s10518-022-01526-5>
1826
1827 Weatherill G, Kotha SR, Danciu L, et al (2024) Modelling seismic ground motion
1828 and its uncertainty in different tectonic contexts: challenges and application
1829 to the 2020 European Seismic Hazard Model (ESHM20). *Natural Hazards and*
1830 *Earth System Sciences* 24(5):1795–1834. <https://doi.org/https://doi.org/10.5194/nhess-24-1795-2024>
1831
1832 Zaccarelli R, Bindi D, Strollo A, et al (2019) Stream2segment: An Open-Source
1833 Tool for Downloading, Preprocessing, and Visually Inspecting Seismic Waveform
1834 Datasets. *Seismological Research Letters* 90(5):2028–2038. <https://doi.org/10.1785/0220180314>
1835
1836
1837
1838
1839
1840

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Georgesetal26suppfiles.zip](#)
- [SupplementaryfiguresGeorgesetal26.pdf](#)