Soil Dynamics and Earthquake Engineering 197 (2025) 109521

Contents lists available at ScienceDirect

Soil Dynamics and Earthquake Engineering

journal homepage: www.elsevier.com/locate/soildyn

Ground motion model using simulated scenario earthquake records in Azores Plateau (Portugal) at bedrock

Kun Ji [a], Shaghayegh Karimzadeh [b,*], Saman Yaghmaei-Sabegh [c], Ruibin Hou [d], Alexandra Carvalho [e], Paulo B. Lourenço [b]

[a] College of Civil and Transportation Engineering, Hohai University, Nanjing, China
[b] Department of Civil Engineering, University of Minho, Institute for Sustainability and Innovation in Structural Engineering (ISISE), ARISE, Guimarães, Portugal
[c] Department of Civil Engineering, University of Tabriz, Tabriz, Iran
[d] School of Architecture and Civil Engineering, Xihua University, 999# Jin Zhou Rd. Jin niu, District, Chengdu, China
[e] National Laboratory for Civil Engineering (LNEC), Lisbon, Portugal

ARTICLE INFO

Keywords:
Ground motion model (GMM)
Stochastic finite-fault approach
Simulated records
Bedrock
Azores Plateau (Portugal)

ABSTRACT

The Azores archipelago in Portugal, located within a seismically active region, experienced several moderate to strong earthquakes throughout its history, including significant events in 1980 (moment magnitude, M_w 6.9) and 1998 (M_w 6.2). Despite its moderate to high seismicity, the region lacks a comprehensive database of recorded ground motions due to limited instrumental seismic data. To address this gap, ground motion simulation techniques provide alternative region-specific time series for areas with sparse seismic networks or a lack of catastrophic earthquake events. This study develops a region-specific ground motion model (GMM) for the Azores Plateau in Portugal, utilizing a homogeneous dataset of region-specific simulated records that have been generated in the bedrock through a stochastic finite-fault approach. The GMM is constructed using a mixed-effects algorithm to predict peak ground acceleration, peak ground velocity, and spectral acceleration ordinates at periods between 0.02 and 2.0 s. The model utilizes input parameters for prediction, including M_w, Joyner-Boore distance (R_{JB}), and focal depth (F_D). The model is formulated for shallow seismic events ranging from magnitude M_w 5.0 to 6.8, F_D 5–17 km, and R_{JB} up to 150 km on bedrock sites. Uncertainty quantification is performed through residual analysis, offering insights into inter-event and intra-event variabilities. The results demonstrate that the proposed model effectively predicts ground motion parameters across the considered range of magnitudes, distances, and periods, providing a valuable tool for assessing the earthquake hazard in the Azores region.

1. Introduction

Earthquakes are natural hazards that can cause considerable loss of life and economic devastation, especially in areas with moderate to high seismic hazard and high vulnerability assets. Although earthquakes affect a smaller segment of the global population compared to other natural disasters, they account for a substantial portion of total financial losses and remain the leading cause of natural disaster-related fatalities [1]. Over the past 40 years, the number of people exposed to moderate to severe earthquakes has nearly doubled due to the growing population and urban expansion [1].

A comprehensive evaluation of the seismic hazard is essential for effectively managing the risk in earthquake-prone areas. This assessment can be carried out using deterministic or probabilistic methods, as summarized by Baker et al. [2]. Ground motion models (GMMs) are critical tools for predicting earthquake impact and evaluating associated hazards. Numerous GMMs have been developed to estimate ground motion parameters, such as spectral response ordinates. Their applicability to various regions has been examined (e.g., NGA-West2 global GMMs summarized in Refs. [3–9]). To accurately evaluate seismic hazard, these models incorporate several key parameters, such as earthquake magnitude, fault mechanism, source-to-site distance, and local site conditions.

As summarized by Douglas [10], GMMs can generally be classified into empirical and simulation-based models. Additionally, some nonparametric GMMs are constructed using machine learning (ML)

* Corresponding author.
E-mail address: shaghkn@civil.uminho.pt (S. Karimzadeh).

https://doi.org/10.1016/j.soildyn.2025.109521
Received 27 February 2025; Received in revised form 3 May 2025; Accepted 7 May 2025
Available online 23 May 2025
0267-7261/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

techniques, such as artificial neural networks, random forests, and gradient boosting [11]. Choosing between these approaches depends on the specific requirements of the hazard assessment and the available data.

GMMs in the literature are developed from either a global dataset (like NGA-West2 database) or a dataset tailored to a specific region's seismological characteristics [12–19]. When there is sufficient data from seismic networks, recorded ground motion data can be used to create these GMMs. However, in cases where data is limited, especially when developing region-specific models, ground motion records can be generated through ground motion simulation techniques [20], including both source-based and site-based approaches [17,18,21-23].

Despite considerable advancements, developing region-specific GMMs remains challenging in many areas, as high-quality data for large-magnitude or near-field seismic events are often scarce. This scarcity hinders the creation of consistent ground motion datasets that accurately represent local seismological characteristics and hazard levels. The Azores Plateau in Portugal is one such region facing this data limitation. Although it is an area of moderate to high seismicity, recorded ground motion data in the Azores are limited, making it difficult to establish accurate, region-specific seismic hazard assessment and risk mitigation models. Previous research has concentrated on constructing an ML-based backbone GMM for the Azores Plateau by developing a homogeneous simulated dataset [18]. Building on this foundation, the current study aims to develop a GMM tailored to the Azores Plateau. To overcome the lack of recorded ground motion data, the study utilizes the stochastically simulated dataset outlined by Karimzadeh et al. [18]. This approach has facilitated the generation of a comprehensive dataset that closely aligns with the regional seismicity and ground motion patterns. This parametric model offers the advantage of clear, interpretable formulas, which are more suitable for classical hazard assessment approaches compared to complex ML models. This transparency allows for more straightforward implementation and verification in practical applications, providing a reliable basis for seismic hazard assessment.

This study utilizes the mixed-effects algorithm by Abrahamson and Youngs [24] and incorporates the modified analytical likelihood functions proposed by Mohammadi et al. [15]. To construct the GMM, the study uses a functional form including both source-scaling and path-scaling terms. Simulations by Karimzadeh et al. [18] were conducted for various scenarios with magnitudes ranging from M_w 5.0 to 6.8, using a bin size of 0.1. These simulations were performed on the EXSIM12 platform [25] through the stochastic finite-fault approach [26,27] by incorporating region-specific input parameters and managing uncertainties in fault ruptures through Monte Carlo methods. The simulation-based GMM utilizes moment magnitude (M_w), Joyner-Boore distance (R_{JB}), and focal depth (F_D) as the primary input parameters. The model produces peak ground acceleration (PGA), peak ground velocity (PGV), and 5 % damped pseudo-spectral acceleration (PSA) within a period range of 0.02–2.0 s as outputs. The performance of the developed GMM herein is assessed using metrics such as coefficient of determination (R^2), Pearson correlation coefficient (r), mean absolute percentage error (*MAPE*), and root mean square error (*RMSE*). Next, a normality test is conducted on the inter-event and intra-event residuals to ensure their unbiasedness. Finally, the developed model is compared against empirical GMMs from the literature, including those by Atkinson (2010) [28], Bradley (2013) [29], Boore et al. (1997). [30], Kotha et al. [6], and one of the NGA-West2 GMMs [31] (proposed by Boore et al., in 2014), as well as the local nonparametric model proposed by Karimzadeh et al. [18] for the region. The results demonstrate the proposed model's ability to effectively capture the complex behavior of seismic ground motions.

2. Ground motion dataset

The central and eastern Azores islands in Portugal – Faial, Pico, São Jorge, Terceira, and Graciosa – each display unique tectonic characteristics (Fig. 1). The fault pattern in the region consists of two main systems with opposing dips, revealing maximum horizontal tensile stress oriented NE-SW, horizontal compressive stress NW-SE, and vertical intermediate compressive stress. Secondary stress fields in eastern São Miguel and Graciosa island may alternate with the primary field [32,33]. The island morphology reflects the interaction between volcanic activity, faulting, and denudation processes, affecting formations from the Middle Pleistocene to the Holocene. Fig. 1 illustrates the active tectonics in the central and eastern Azores islands. Information regarding the active faults is provided in Table 1. According to Madeira et al. [34], the maximum expected magnitude for various faults is determined using the empirical equation suggested by Wells and Coppersmith [35].

A previous study by Karimzadeh et al. [18] conducted extensive simulations that generated a comprehensive dataset of 247,710 ground motion records across the entire Azores Plateau. These authors performed simulations using the EXSIM12 platform [36,37] to model the acceleration time series of various scenario events. The latest version of the stochastic finite-fault ground motion simulation approach is applied to generate acceleration time series of scenario earthquakes [37]. The methodology, originally based on the FINSIM code developed by Beresnev and Atkinson [38], has been enhanced through the improvements proposed by Boore [39]. The enhanced version of this approach improves the low-frequency component of the simulations. By considering factors such as earthquake magnitude, fault geometry, strike, dip, slip distribution, density, and rupture velocity, the method accurately simulates the fault rupture.

The study by Karimzadeh et al. [18] considered 23 scenario events with magnitudes ranging from M_w5.0 to 6.8, in 0.1 magnitude increments, corresponding to the rupture of different active onshore faults in the region (see red fault traces in Fig. 1). Among these, nine are on Faial Island (F1, F2E, F2W, F3, F4, F5, F6, F7, F8, as detailed in Table 1), and one on São Jorge Island (SJ1, as detailed in Table 1). The red fault traces in Fig. 1 represent the faults in the region used for simulations in Karimzadeh et al. [18], while the black traces indicate faults with similar tectonic characteristics. Among these representative active faults, each exhibiting distinct characteristics, four are associated with a maximum magnitude of M_w 6.3 and two with M_w 5.2, while a single scenario represents each of the remaining magnitudes. This uneven distribution explains the relatively larger number of simulated recordings for M_w 5.2 and M_w 6.3 in our analysis. We focused solely on the rupturing of nearby onshore fault models, as suggested by the study of Madeira et al. [34]. According to Madeira et al. [34], the fault length suggests a maximum expected magnitude of 6.8 within the region using the empirical equation suggested by Wells and Coppersmith [35]. The authors propose a maximum expected magnitude of M_w 6.6 for Faial Island, but this value increases to M_w 6.8 when considering the potential rupture of the Picos fault in São Jorge Island. Thus, the maximum magnitude modeled in our study is M_w 6.8. These simulations were not confined to a single fault but covered various active onshore faults in the Azores Plateau, accounting for aleatory region-specific uncertainty using the established framework.

The simulations were completed across 359 dummy stations (see Fig. 1) on all islands with a grid size of 1 km and were performed on bedrock. Region-specific input parameters provided by studies [40,41,18] were utilized for calibration based on validations against observed motions from the 1998 Faial earthquake (M_w 6.2). To account for uncertainty in source and attenuation effects, key parameters such as hypocenter location, stress drop, pulsing percent, quality factor, and kappa [26,39] were treated as random variables. The regional model for input parameters with probability distribution functions (PDFs) and their ranges was derived from Carvalho et al. [42] (see Table 2). It is noted that for each event, 30 Monte Carlo simulations were carried out, each with distinct combinations of input parameters. As a result, each scenario event was simulated with 30 different sets of random input-model

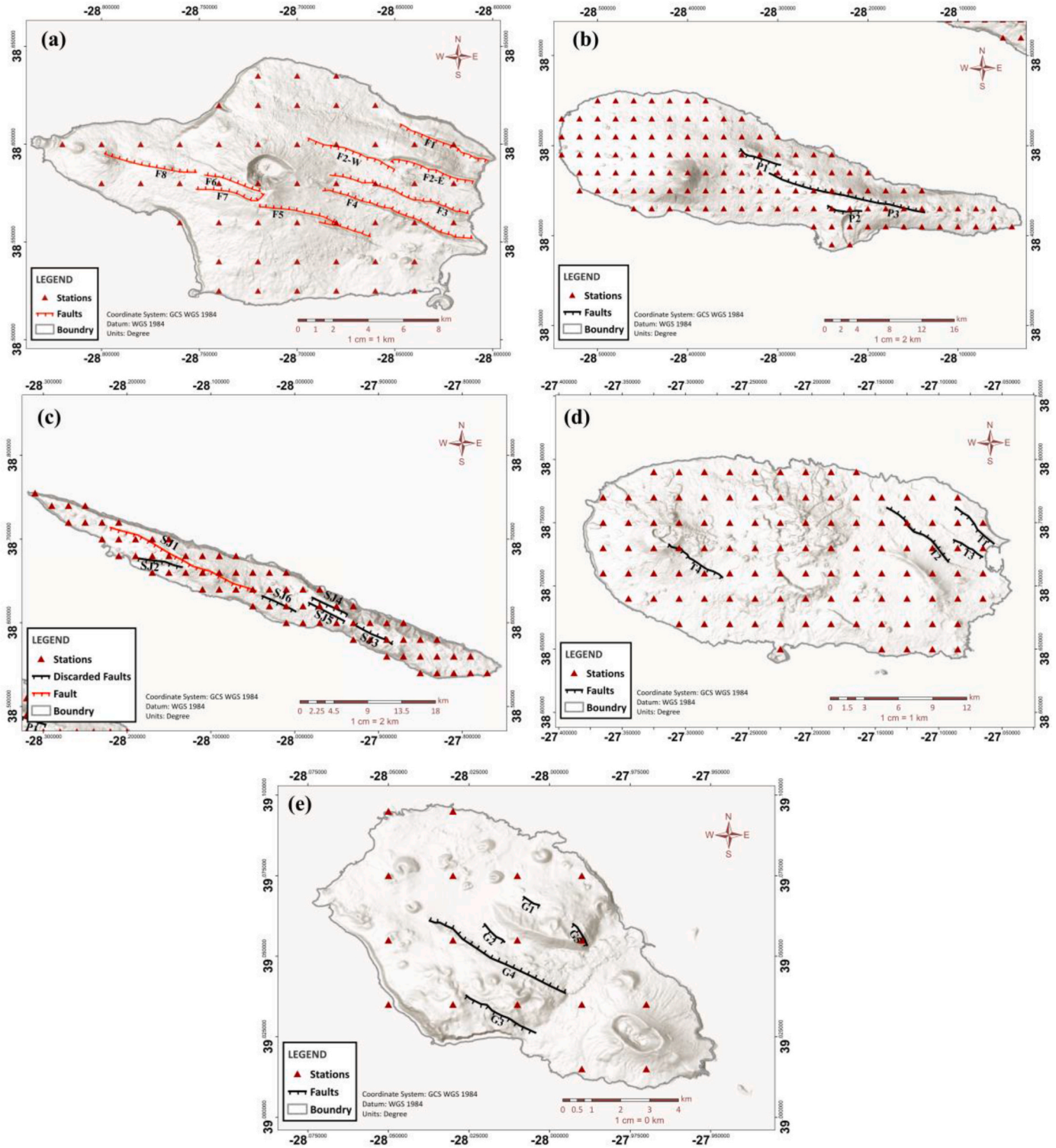


Fig. 1. Active faults on the central and eastern Azores islands (a) Faial, (b) Pico, (c) São Jorge, (d) Terceira, and (e) Graciosa. The dummy stations are shown by triangles. The red fault traces represent the faults in the region used for simulations in Karimzadeh et al. [18], while the black traces indicate faults with similar tectonic characteristics. The 359 dummy stations on the islands are represented by red triangles. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

parameters, yielding 247,710 simulations across 359 sites. It is also noted that the simulated database was generated from ten fault ruptures, nine on Faial Island and one on São Jorge Island. More detailed information can be found in the study by Karimzadeh et al. [18]. It is important to highlight that these simulations have additionally been validated and explored in the studies by Karimzadeh et al. [43] and Bernardo et al. [44], focusing on the seismic demand evaluation of masonry cultural heritage. This expansive dataset is the foundation for the current analysis, providing a robust data source for evaluating ground motions in the region. The data have been meticulously processed using baseline correction and a 4th-order Butterworth filter, with a frequency range from 0.1 to 25 Hz. These measures ensure data

Table 1
Information on the active faults for the central and eastern Azores islands [18].

Region	No	Fault Name	Fault Rupture Length (km)	$M_{w\text{-max}}$	Fault Mechanism	Strike (°)	Dip (°)
Faial	F1	Ribeirinha	12.5	6.3	Normal	115	75
Faial	F2-E	Lomba Grande Eastern segment	12.5	6.3	Normal	115	80
Faial	F2-W	Lomba Grande Western segment	12.5	6.3	Normal	115	80
Faial	F3	Rocha Vermelha	14	6.4	Normal	290	55
Faial	F4	Espalamaca	20.3	6.6	Normal	295	70
Faial	F5	Flamengos	11.5	6.3	Normal	290	70
Faial	F6	Lomba do Meio	4	5.2	Normal	295	70
Faial	F7	Lomba de Baixo	4	5.2	Normal	300	50
Faial	F8	Capelo	8.8	5.8	Normal	290	90
Graciosa	G1	Saúde-Hortelã	5	5.9	Normal	140	–
Graciosa	G2	South Serra das Fontes	4.6	5.8	Normal	126	–
Graciosa	G3	North Serra Branca	4.8	5.9	Normal	302	–
Graciosa	G4	South Serra Branca	3.2	5.7	Normal	305	–
Graciosa	G5	East Serra das Fontes	4.6	5.8	Normal	340	–
Pico	P1	Lagoa do Capitão	8.8	6.2	Normal	120	80–90
Pico	P2	Topo	7.5	6.1	Normal	285	70–90
Pico	P3	Cabeço do Sintrão	21	6.6	Normal	293	–
São Jorge	SJ1	Picos	33	6.8	Normal	120	90
São Jorge	SJ2	Pico Carvão	12	6.3	Normal	285	75–90
São Jorge	SJ3	Urze-São João	15	6.4	Normal	304	80
São Jorge	SJ4	Cume Faja do Belo	7.4	6.1	Normal	120	70
São Jorge	SJ5	Serra do Topo	7.2	6.1	Normal	140	–
São Jorge	SJ6	Ribeira Seca	7.3	6.1	Normal	160–170	–
Terceira	T1	Lajes	8.2	6.1	Normal	138	70–90
Terceira	T2	Fontinhas	9	6.2	Normal	313	–
Terceira	T3	Cruz do Marco	5	5.9	Normal	310	70
Terceira	T4	Santa Bárbara	12.9	6.4	Normal	308	70

Table 2
Input-model parameters of simulations [18].

Parameter	Value
Crustal Thickness, D (km)	13
Crustal Density (g/cm^3)	$Depth = 0.0\ km \rightarrow 2.67$ $Depth = 2.5\ km \rightarrow 2.77$ $Depth = 8.0\ km \rightarrow 2.86$ $Depth = 14.0\ km \rightarrow 2.93$
Shear Wave Velocity (km/s)	$Depth = 0.0\ km \rightarrow 3.1$ $Depth = 2.5\ km \rightarrow 3.7$ $Depth = 8.0\ km \rightarrow 4.2$ $Depth = 14.0\ km \rightarrow 4.6$
Shear Wave Velocity/Crustal Velocity	0.8
Geometric Spreading	$R^{-1.0}$ $R \leq 1.5D\ km$ $R^{0.0}$ $1.5D\ km < R \leq 2.5D\ km$ $R^{-0.5}$ $R > 2.5D\ km$
Duration Model (R in km)	T_0+0.1R
Window Type	Saragoni-Hart
Damping	5 %
Slip Weight	Random
Iseed	309
Hypocentre Location	Uniform distribution: along the length and width
Pulsing Percent	Uniform distribution: 30-50
Kappa	Uniform distribution: 0.075 ± 0.02
Stress Drop (bars)	Lognormal distribution: 110 ± 20
Quality Factor	Lognormal distribution: $(76 \pm 11)f^{0.69\pm0.09}$

consistency and high quality, minimizing noise and other anomalies that could affect the accuracy of our models.

Fixing kappa values were derived from rock or very stiff soil records (V_{S30}> 800 m/s) in the region, as established in the study by Carvalho et al. [42]. Although minor amplification effects may exist, we acknowledge the limited availability of records explicitly measured on bedrock. Carvalho et al. [42] explicitly note that "site amplifications were neglected, as records analyzed were obtained at sites classified as rock or very stiff ground." Therefore, our fitted GMM is based on bedrock conditions without accounting for site amplification effects. Our proposed GMM is not explicitly tied to a specific site class or V_{S30} value. Different building codes and seismic design specifications adopt varying definitions for reference bedrock conditions (i.e., unamplified rock sites), with the reference V_{S30} thresholds ranging from 760 m/s to over 1000 m/s. Given this variability, our model is intended for application to unamplified or minimally amplified rock sites, consistent with widely accepted definitions of reference rock used in seismic design.

This dataset is rich in seismological characteristics pertinent to the Azores region, including a homogeneous distribution in terms of M_w, R_{JB}, and F_D. The dataset also encompasses R_{JB} distances ranging from 0 to 150 km, offering a detailed representation of both near-field and far-field seismic data. The data at approximately R_{JB} 90 km is unavailable due to the presence of islands, as these distances fall offshore. FD values range from 5.0 to 17.0 km, indicating the dominance of shallow events within the dataset. This breadth of data provides a nuanced understanding of the seismological features specific to the Azores region. Fig. 2 provides a comprehensive overview of all statistical details pertaining to the dataset, offering insight into its key characteristics.

The minimum magnitude threshold of M_w 5.0 in our analysis reflects the lower bound derived from the simulated ground motion dataset for active faults in the study area [18]. This threshold was established considering the practical limitations of the stochastic finite-fault simulation approach, which is less effective for modeling small-magnitude events due to their short rupture dimensions [26]. However, numerous studies have reported structural damage caused by small-magnitude earthquakes worldwide, as well as the notable contribution of such events to seismic hazard, particularly when generated by complex faulting systems (e.g., Ref. [45]). Specifically, probabilistic seismic hazard analyses may underestimate the annual exceedance probability of short-period ground motions if lower magnitude events are not considered, particularly in high-seismicity regions with clustered low-magnitude activity.

The proposed GMM is developed based on simulated data with moment magnitudes up to M_w 6.8, which corresponds to the largest expected earthquake in the region of interest, as given in Madeira et al. [34]. However, in probabilistic seismic hazard analysis, it is common practice to consider scenario events with magnitudes slightly exceeding the historical maximum (e.g., M_{max} + 0.1 or M_{max} + 0.2) to account for epistemic uncertainties in the characterization of seismic sources. Since the current GMM has not been trained or validated for magnitudes beyond M_w 6.8, its applicability to such scenarios remains uncertain.

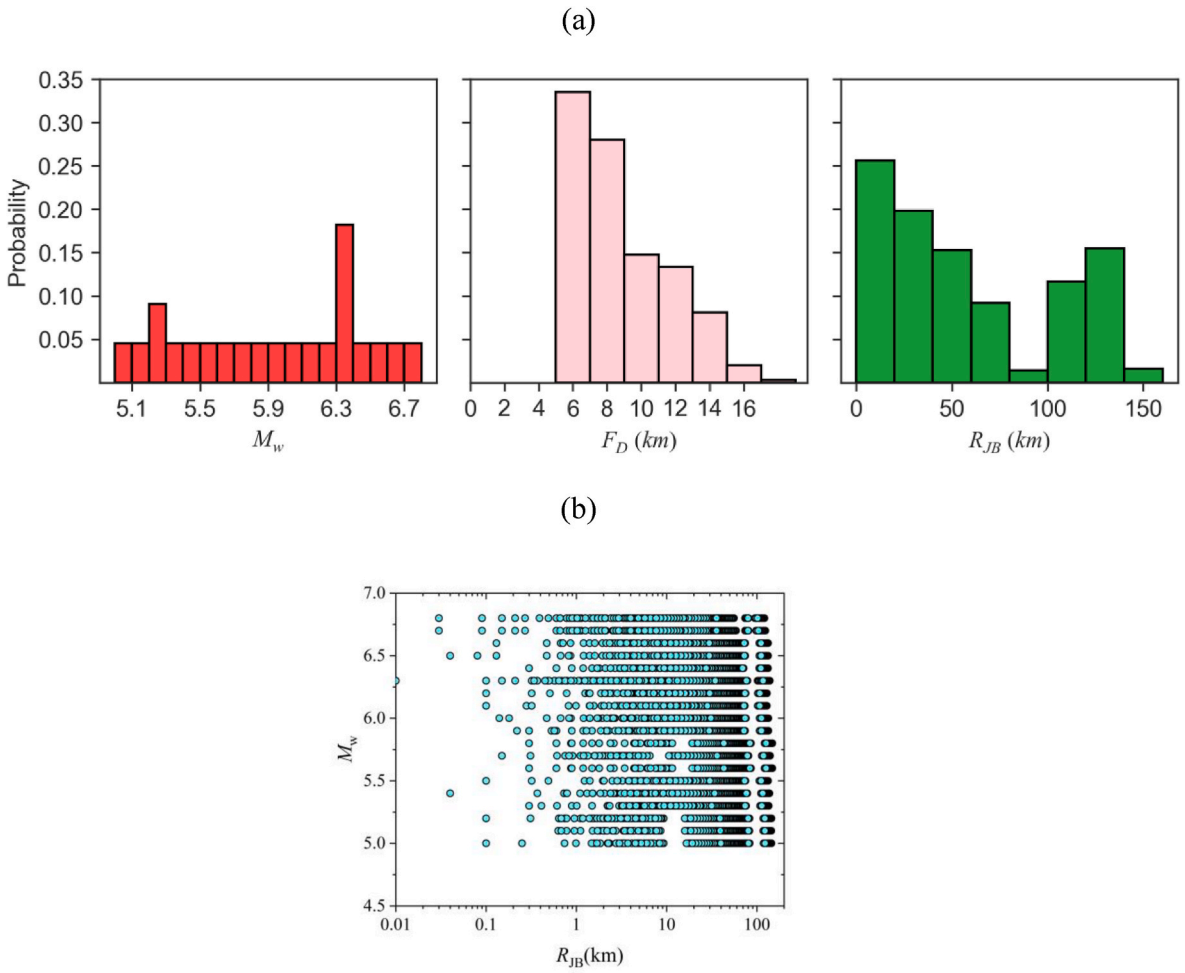


Fig. 2. (a) Seismological properties of the Azores ground motion dataset in terms of moment magnitude (M_w), Joyner and Boore distance (R_{JB}), and focal depth (F_D), (b) Distribution of M_w with respect to R_{JB}.

Direct extrapolation of the model beyond its calibrated magnitude range may introduce significant prediction bias, especially in near-field ground motion characteristics. This limitation may affect the accuracy of hazard estimates if the model is used outside its calibrated magnitude range. Future work may consider extending the synthetic ground motion dataset to include higher magnitudes (e.g., M_w 6.9 or 7.0) to improve the applicability of the model in hazard assessments involving larger events.

3. Functional form for median ground motion model

To build the GMM, it is necessary to select a functional form guided by a visual inspection of the data distribution trend, as shown in Fig. 3, which illustrates the magnitude and distance dependence of different IMs for all 247,710 stochastically simulated ground motions. The data are separated into magnitude bins and presented in panels for various ground motion intensity measures (IMs). Although the recordings are simulated through a stochastic approach, the magnitude and distance dependence of different IMs in our simulated results are similar to the data distribution of real recordings in the literature (e.g., the NGA-West2 dataset in Boore et al. [31]). Close inspection of the data-amplitude plots in Fig. 3 reveals several key features that the functional form must accommodate: magnitude-dependent geometric spreading, anelastic attenuation effects evident from the curvature in the decay of IM amplitude versus ln(R_{JB}) in the far field, and a tendency for magnitude saturation to be observed with increasing magnitude for PSA at short periods and close distances, indicating a strongly nonlinear (and period-dependent) magnitude dependence of amplitude scaling at a fixed distance. The saturation is less pronounced toward longer periods and longer distances, as seen in PSA ($T = 1.5$ s) in Fig. 3.

The median prediction of the GMM is given by the general equation:

$$\ln \widehat{Y} = f_{\text{mag}} + f_{\text{dis}} \tag{1}$$

where $\ln \widehat{Y}$ represents the median prediction of the ground motion IM (PGA, PGV, or PSA ranging from 0.02 to 2.0 s) in logarithm. Stochastic approaches are limited in accurately capturing long-period ground motions and tend to exhibit higher uncertainty at low frequencies. Consequently, our GMM is constrained to a maximum period of 2.0 s, the threshold for acceptable accuracy as identified by Karimzadeh et al. [18]. The source-term and the path-term, f_{mag} and f_{dis}, are formulated as functions of M_{w}, R_{JB}, and F_{D}. The source-term f_{mag} is given by:

$$f_{\text{mag}} = \begin{cases} a_0 + a_1 M_{\text{w}}; & M_{\text{w}} < 5.5 \\ a_0 + a_1 M_{\text{w}} + a_2 (M_{\text{w}} - 5.5); & 5.5 \le M_{\text{w}} < 6.5 \\ a_0 + a_1 M_{\text{w}} + a_2 (M_{\text{w}} - 5.5) + a_3 (M_{\text{w}} - 6.5); & M_{\text{w}} \ge 6.5 \end{cases} \tag{2}$$

The path-term f_{dis} contains two components, magnitude-dependent geometric spreading and apparent anelastic attenuation:

$$f_{\text{dis}} = \underbrace{\left[a_4 + a_5 \times (M_w - M_{ref})\right] \ln \sqrt{{R_{\text{JB}}}^2 + h^2}}_{\textit{geometric spreading}} + \underbrace{a_6 \left(\sqrt{{R_{\text{JB}}}^2 + {F_{\text{D}}}^2} \right)}_{\textit{apparent anelastic attenuation}} \tag{3}$$

The terms a_0, a_2, …, a_6 denote the coefficients of the regression model. In the stochastic simulation process, only the normal fault type is

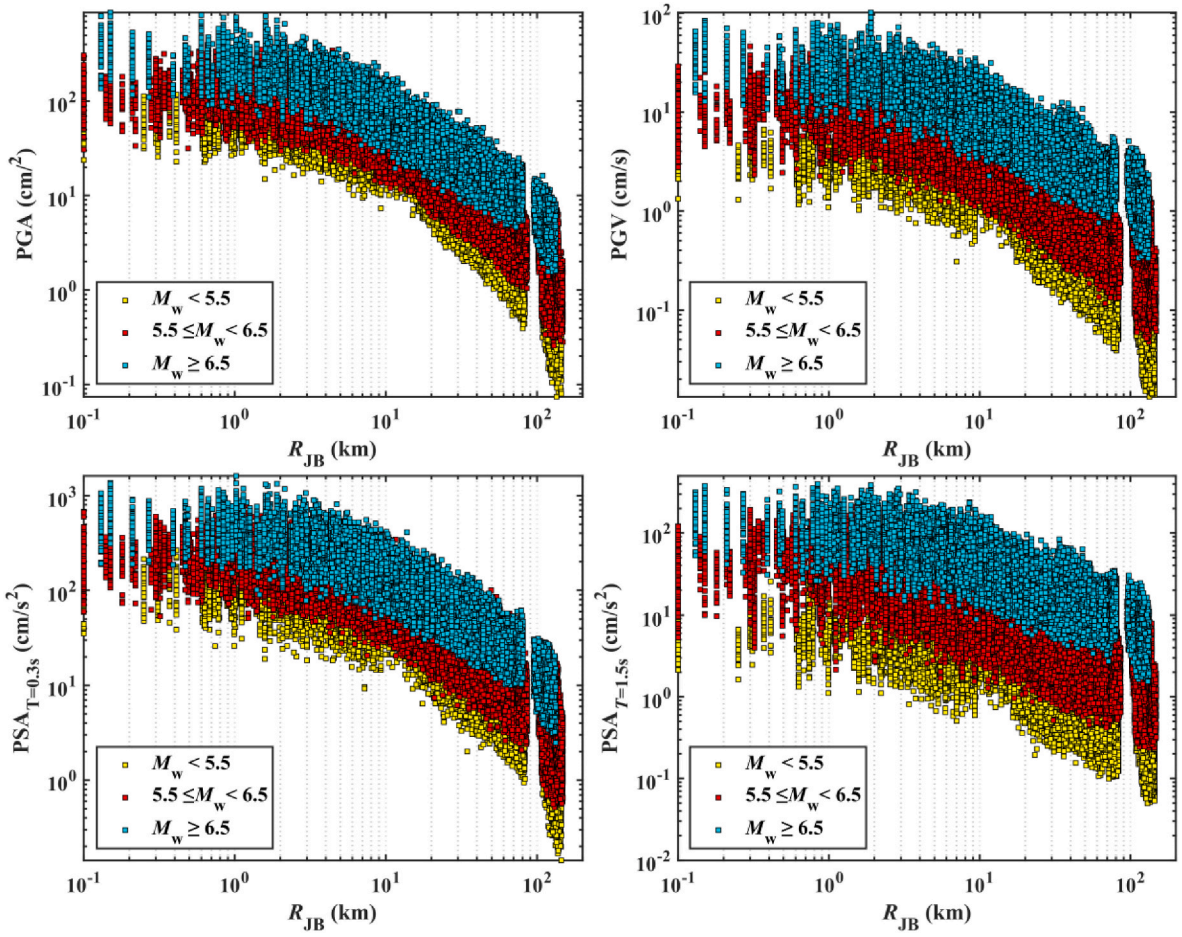


Fig. 3. PGA, PGV, $PSA_{T\ =\ 0.3s}$, and $PSA_{T\ =\ 1.5s}$ versus R_{JB} for various magnitude ranges.

considered based on the information of the faults listed in Table 1. As a result, the style-of-faulting term is not included in our GMM. Hanging wall effects are not considered because the distance is measured by $R_{JB,}$ which implicitly accounts for larger motions over the hanging wall [46].

As noted before, site amplification factors are not included as input parameters in the stochastic finite-fault simulation framework. Instead, amplification is assumed to be one (i.e., no amplification), and therefore, our fitted model is intended to represent theoretical bedrock conditions without accounting for site response effects. Next, when comparing with traditional empirical GMMs, we assume a site amplification factor of one by adopting the reference V_{S30} value defined in each model (e.g., 760 m/s in Boore et al. [31]), thereby ensuring consistency in ground motion predictions under equivalent rock site conditions. It is worth noting that the reference model for different empirical GMMs is regressed under various hinge V_{S30}. The site amplification factor is expected to be one for sites with the hinge V_{S30}, although the so-called bedrock site condition still has some amplification compared to the simulation results. In future studies, site amplification should be incorporated using various methods, such as empirical or semi-empirical regional site amplification models derived from a combination of geology and slope [47,48,49], code-based site amplification factors [50], or horizontal-to-vertical spectral ratio (HVSR) [51], 1D or 2D site response analyses [52], or theoretical models, to provide a more comprehensive framework for site-specific ground motion prediction and application.

In the source-term f_{mag}, a piecewise trilinear functional form is utilized instead of the quadratic function. This is because the trilinear function form for magnitude scaling performs better than the quadratic function form according to the Akaike information criterion (AIC) and Bayesian information criterion (BIC) calculation results. Because the utilized earthquake dataset ranges from M_w 5.0 to 6.8, three linear scaling terms are sufficient to represent the magnitude saturation effect without the need to consider small-magnitude recordings. Breakpoints are set at $\boldsymbol{M_w}$ 5.5 and 6.5 in the magnitude scaling term, as shown in Eq. (2) and Fig. 3. In a similar study by the NGA-West2 model [53], a trilinear function form is also used for magnitude scaling, with breakpoints set as M_w 4.5, 5.5, and 6.5. Considering that the magnitude range for our dataset is M_w 5.0 to 6.8, choosing M_w5.5 and 6.5 as breakpoints in magnitude scaling is reasonable. We also tested other hinge magnitude breakpoints, such as M_w6.0 and 6.5, which yielded higher maximum likelihood values and lower corresponding AIC and BIC values.

For the path-term f_{dis}, it is necessary to constrain the pseudo-depth h in the regression to prevent overlap in the curves for large earthquakes at very short distances. Since the ground motion dataset only includes shallow-to-intermediate depth events ranging from 5 to 17 km, there is no need to divide the focal depth into groups. Therefore, the pseudo-depth h can be treated as depth-independent and assigned a priori value based on nonlinear regression trials as in previous studies (Kotha et al. [54,6]). Following nonlinear regression trials for various IMs, h is set as 3.0 km, which is comparable with other similar studies. In the NGA-West2 GMM proposed by Boore et al. [31], h ranges from 4.04 to 9.65 km across various IMs. In the GMM proposed by Kotha et al. [6], h is set to 4.0 km for shallow events with $F_D < 10$ km and 8.0 km for events of intermediate depth 10 km $\leq F_D <$ 20 km. We could also perform initial regressions with h as a free parameter, then adjust the obtained values of h as needed to avoid spectral overlap at close distances. However, this process is inefficient and does not significantly improve performance based on our trials. After pseudo-depth h is determined, the coefficients

a_4, a_5, and a_6 are obtained through regression analysis.

4. Computational efficient regression method adapted to large datasets

Our mixed-effects GMM is composed of fixed effects and random effects, as shown in Eq. (4):

$$\ln\left(y_{ij}\right) = \ln\widehat{y} + \eta_i + \varepsilon_{ij} \tag{4}$$

where the index i signifies the earthquake event, and j indicates the station index. The term y_{ij} represents the IM for the i^{th} event at the j^{th} station. $\widehat{y}$ is the median prediction, which represents the fixed effects without any specificities related to the event and site. The term η_i represents the inter-event residual component while ε_{ij} denotes the intra-event residual component in the natural logarithm scale. In GMMs, these two types of residuals, namely inter-event and intra-event residuals, are assumed to be independent, normally distributed random variables with a mean of zero and standard deviations of τ and ϕ, respectively. Given the independence of inter-event and intra-event residuals, the total standard deviation (σ) is partitioned into components that represent between-event variability (τ) and within-event variability (ϕ), as follows:

$$\sigma = \sqrt{\tau^2 + \phi^2} \tag{5}$$

The widely recognized mixed-effects model approach proposed by Abrahamson and Youngs [24] is frequently used in the literature for conducting residual analysis. As suggested by these authors, the maximum likely solution for the random effect, inter-event residual for each earthquake is defined as follows:

$$\eta_i = \frac{\tau^2 \sum_{j=1}^{n_i}\left(y_{ij} - \widehat{y}_{ij}\right)}{n_i\tau^2 + \phi^2} \tag{6}$$

where $\widehat{y}_{ij}$ is the median prediction value using the proposed GMM as shown in Eq. (1).

According to Abrahamson and Youngs [24], the primary likelihood function is defined as follows:

$$\ln L = -\frac{N}{2}\ln 2\pi - \frac{1}{2}\ln|\mathbf{C}| - \frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})^T\mathbf{C}^{-1}(\mathbf{y}-\boldsymbol{\mu}) \tag{7}$$

where N is the number of data points, $\mathbf{C}$ is the covariance matrix, $\boldsymbol{\mu}$ is the vector of predicted values, and $\mathbf{y}$ is the vector of observed values.

Mohammadi et al. [15] modified the procedure by incorporating an algebraic maximum likelihood function to estimate model parameters and variances using the expectation-maximization algorithm. The modified formula for the maximum likelihood function is defined as:

$$\ln L = -\frac{N}{2}\ln 2\pi - \frac{N-M}{2}\ln\phi^2 - \frac{1}{2}\sum_{i=1}^{M} ln\left(\phi^2 + n_i\tau^2\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{M}\sum_{j=1}^{n_i}\left(y_{ij} - u_{ij}\right)^2 + \frac{r^2}{2\phi^2}\sum_{i=1}^{M}\left(\frac{n_i^2}{\phi^2 + n_i\tau^2}(\overline{Y} - \overline{\mu}_i)^2\right) \tag{8}$$

where N is the total number of records, M is the total number of events, n_i is the number of records for the i^{th} event, and the terms $\overline{Y}_i$ and $\overline{\mu}_i$ are, respectively, the mean values of observed and predicted IM for the i^{th} event. A detailed derivation process is given in the Appendix of Mohammadi et al. [15]. Eq. (8) is both straightforward and computationally efficient, making it particularly useful for deriving GMMs when handling a large number of events, as is the case in this study.

The iterative procedure developing mixed effect GMM model is summarized as follows.

Step 1. The model parameters θ in Eq. (9) are estimated using a fixed effect regression procedure. The term y_{ij} represents the IM for the i^{th} event at the j^{th} station as defined in Eq. (4). The median GMM prediction value $\ln\left(\widehat{y}_{ij}\right)$ is given by the general function as described in Eq. (1). δ_{ij} is the total standard deviation.

$$\ln\left(y_{ij}\right) = \ln\left(\widehat{y}_{ij}\right) + \delta_{ij} = f(M_w, R_{JB}, F_D, \theta) + \delta_{ij} \tag{9}$$

M_{ref} and h are constrained to prior values, and then all the coefficients a_0 to a_6 (Eq. (2) and Eq. (3)) are regressed.

Step 2. Given the estimated θ, the variance ϕ^2 and τ^2 are estimated by maximizing the likelihood function defined in Eq. (8). The solution for minimizing or maximizing a nonlinear unconstrained multivariable objective function is obtained using the derivative-free search method proposed by Lagarias et al. [55] without using numerical or analytic gradients. This method is conveniently implemented using the "fminsearch" function in MATLAB.

Step 3. Based on the estimated values of (σ,τ) and model parameters θ, the random-effect term, inter-event residuals η_i are obtained using the likelihood function from Eq. (6). In this study, given that the number of records per event is substantial (with $n_i = 359$), and that $n_i\tau^2$ greatly exceeds ϕ^2, the approximate equation serves as an accurate method for estimating inter-event residuals. The approximate equation is given as follows:

$$\eta_i \approx \frac{\sum_{j=1}^{n_i}\left(y_{ij} - u_{ij}\right)}{n_i} \tag{10}$$

Step 4. A new model is trained using a fixed-effect regression procedure for $\ln\left(y_{ij}\right) - \eta_j$, that is:

$$\ln\left(y_{ij}\right) - \eta_j = f(M_w, R_{JB}, F_D, \theta) + \delta_{ij} \tag{11}$$

Step 5. Steps 2, 3, and 4 are iterated until the termination criterion is fulfilled. The adopted termination criterion is set as 0.1 % in terms of the difference between two successive likelihood values. Since simulated record datasets provide sufficient data, a relatively strict termination criterion for iterative convergence is adopted.

5. Residual analysis of the developed ground motion model

In Eqs. (4) and (5), Y denotes the IM in terms of PGA, PGV, and PSA ranging from 0.02 to 2.0 s. The terms a_0, a_2, …, a_6 denote the coefficients of the regression model. The regression coefficients, inter-event (τ), intra-event (ϕ), and total (σ) standard deviations for all IMs are calculated using the optimization algorithm as described in the previous section and are presented in Table 3.

Fig. 4 shows the scaling of PGA and $PSA_{T\,=\,1.5s}$ with magnitude for R_{JB} distances of 1, 50, and 110 km. The corresponding simulated data are separated into groups according to the distance range. Because small magnitude events are not included in the development of our GMM, this

Table 3
Regression coefficients and inter-event (τ), intra-event (ϕ), and total (σ) standard deviations for the proposed GMM.

PSA(T)	a_0	a_1	a_2	a_3	a_4	a_5	a_6	ϕ	τ	σ
PGA	−4.967	0.672	−0.277	0.058	−1.072	0.172	−0.010	0.237	0.203	0.312
$T = 0.02$ s	−4.917	0.666	−0.274	0.061	−1.074	0.173	−0.010	0.237	0.204	0.312
$T = 0.03$ s	−4.842	0.656	−0.269	0.064	−1.077	0.173	−0.010	0.236	0.206	0.313
$T = 0.05$ s	−4.453	0.606	−0.246	0.081	−1.094	0.177	−0.011	0.234	0.222	0.323
$T = 0.07$ s	−3.910	0.539	−0.204	0.100	−1.101	0.176	−0.011	0.231	0.248	0.339
$T = 0.10$ s	−3.499	0.499	−0.164	0.125	−1.066	0.163	−0.012	0.234	0.270	0.357
$T = 0.15$ s	−3.673	0.549	−0.156	0.135	−0.980	0.135	−0.013	0.247	0.266	0.363
$T = 0.20$ s	−4.109	0.629	−0.193	0.106	−0.935	0.122	−0.013	0.258	0.247	0.357
$T = 0.25$ s	−4.645	0.718	−0.239	0.065	−0.907	0.113	−0.013	0.267	0.229	0.352
$T = 0.30$ s	−5.283	0.819	−0.301	0.059	−0.880	0.107	−0.012	0.275	0.214	0.348
$T = 0.35$ s	−6.022	0.937	−0.376	0.032	−0.857	0.101	−0.012	0.284	0.201	0.348
$T = 0.40$ s	−6.665	1.040	−0.442	0.015	−0.850	0.098	−0.011	0.291	0.191	0.348
$T = 0.45$ s	−7.282	1.137	−0.509	−0.024	−0.851	0.098	−0.011	0.297	0.183	0.349
$T = 0.50$ s	−7.934	1.240	−0.575	−0.050	−0.847	0.096	−0.010	0.303	0.176	0.350
$T = 0.60$ s	−9.234	1.441	−0.695	−0.132	−0.826	0.092	−0.010	0.315	0.165	0.355
$T = 0.70$ s	−10.413	1.622	−0.793	−0.204	−0.810	0.087	−0.009	0.325	0.156	0.360
$T = 0.80$ s	−11.445	1.779	−0.877	−0.279	−0.804	0.088	−0.009	0.332	0.149	0.364
$T = 1.00$ s	−13.236	2.044	−0.986	−0.420	−0.789	0.084	−0.008	0.346	0.136	0.372
$T = 1.20$ s	−14.614	2.237	−1.032	−0.563	−0.777	0.081	−0.007	0.356	0.124	0.377
$T = 1.50$ s	−16.080	2.428	−1.044	−0.725	−0.772	0.084	−0.006	0.368	0.111	0.385
$T = 2.00$ s	−17.507	2.578	−0.956	−0.873	−0.778	0.092	−0.005	0.385	0.104	0.398
PGV	−3.320	1.127	−0.374	−0.113	−1.080	0.233	−0.007	0.276	0.123	0.302

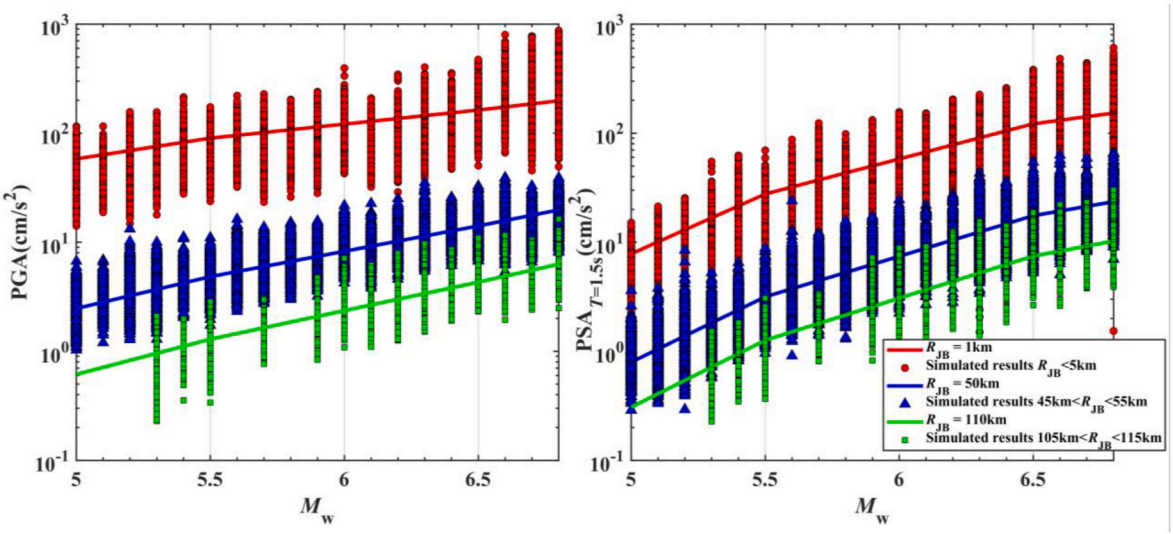


Fig. 4. Scaling of PGA and $PSA_{T\ =\ 1.5s}$ with magnitude for R_{JB} distances of 1, 50, and 110 km.

study specifically examines the magnitude scaling in the range of M_w 5.0 to 6.8. As shown in Fig. 4, the magnitude scaling of PGA is more gradual (less steep) at near-source distances ($R_{JB} = 1$ km) compared with larger distances ($R_{JB} = 100$ km). The slope of magnitude scaling for PGA is overall more gradual than that for PSA at a long period ($PSA_{T\ =\ 1.5s}$), which is similar to previous studies (e.g., NGA-West2 GMM proposed by Campbell and Bozorgnia [53]). Due to a large number of near-source ground motions compiled in our simulation-based dataset, the magnitude saturation effect is already well reflected and constrained in our GMM. Since the maximum considered magnitude in our dataset is M_W 6.8, there is no need to consider PGA oversaturation at large magnitudes ($M_W \geq 6.75$) as discussed in other studies (e.g., Ref. [31,6,54]).

The predicted scaling of PGA and $PSA_{T=1.5\ s}$ with R_{JB} is illustrated in Fig. 5. The near-source saturation for shallow and intermediate depth events is essentially the same, plateauing between 0 and 5 km. The three curves corresponding to F_D of 5, 10, and 15 km converge at approximately 5–10 km. After the merging point, the depth-dependence of distance scaling becomes negligible. Additionally, Fig. 5(c) illustrates the geometrical spreading term, given by $a_4 + a_5$ ($M_w - M_{ref}$). The coefficients exhibit minimal variation with period, except for a slight fluctuation at around 0.4 s. Note that the signs of the a_4 and a_5 coefficients differ for periods less than 2.0 s. As a result, the geometrical spreading factor decreases with magnitude for periods of less than about 2.0 s. In the path-effect term, f_{dis} incorporates magnitude-dependent apparent attenuation through the model coefficient a_6. As expected, the a_4+a_5(M_w-M_{ref}) terms are all negative, indicating attenuation with distance, with the absolute value decreasing (i.e., less decay) as magnitude increases. Regarding the apparent anelastic attenuation coefficient (a_6), this coefficient is well constrained, and the overall trend with periods is similar to previous studies using similar functional form representing anelastic attenuation effect (e.g., c3 value in GMM fitted by Boore et al. [31]). At periods greater than 0.1s and PGV, the apparent anelastic attenuation term of our proposed model is significantly lower (more negative) than that in the GMM of Boore et al. [31], indicating more rapid attenuation in the Azores than the global region considered in BSSA14. The significant anelastic attenuation is related to the low Q_0 value in the volcanic region of the Azores islands (Table 2), which will be further discussed in the comparison with existing GMMs.

Fig. 6 shows the median predicted PSA versus various periods for M_w 5.0, 5.5, 6.0, and 6.8 earthquake scenarios with $R_{JB} = 10$ and 100 km,

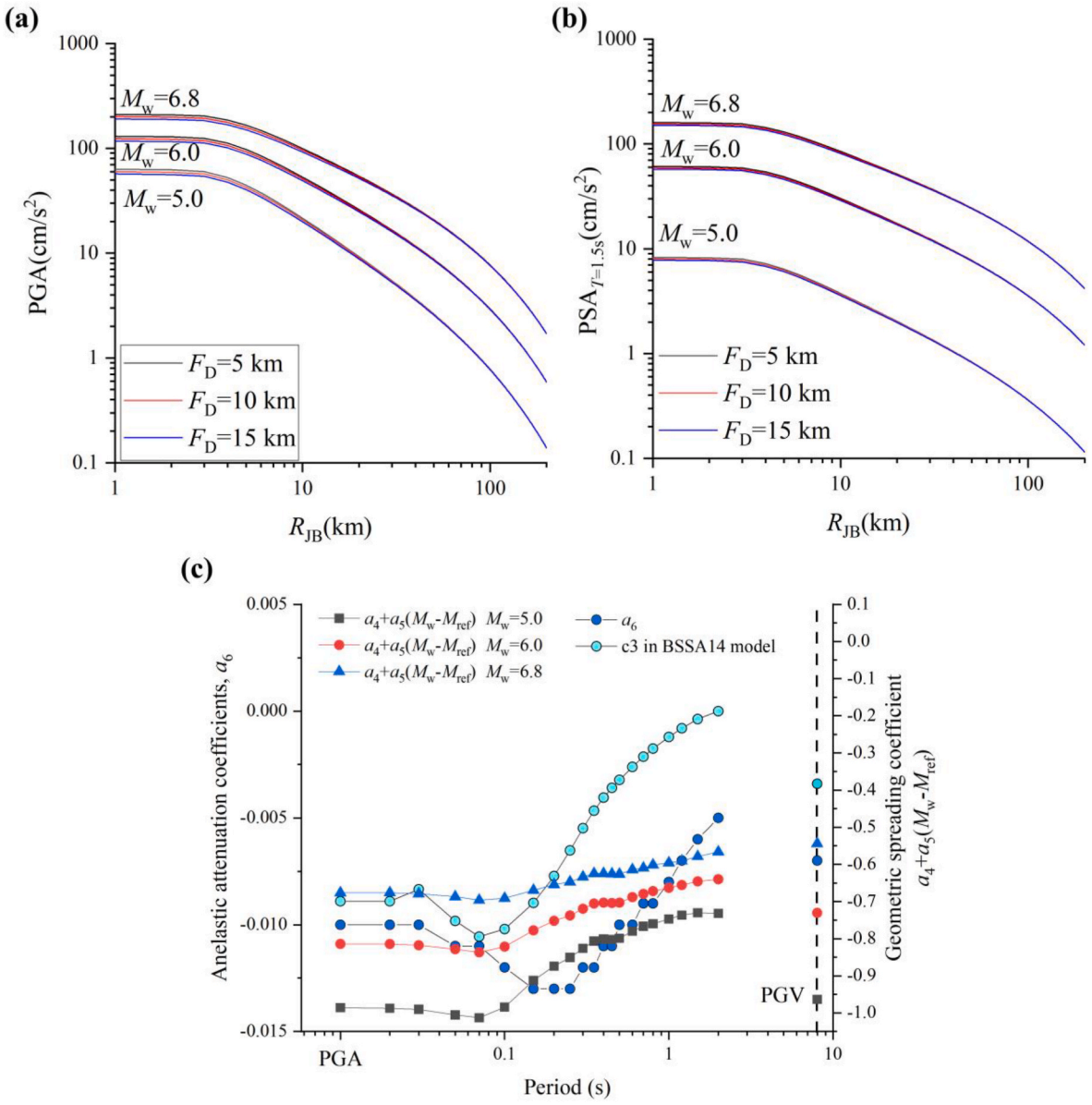


Fig. 5. GMM median prediction for (a) PGA and (b) PSA$_{T\,=\,1.5s}$ for focal depths of 5, 10, and 15 km. (c) effective geometrical spreading coefficient, given by $a_4 + a_5$ ($M_w - M_{ref}$); apparent anelastic attenuation coefficient (a_6) used in our proposed GMM and the corresponding term c3 in BSSA14 GMM [31].

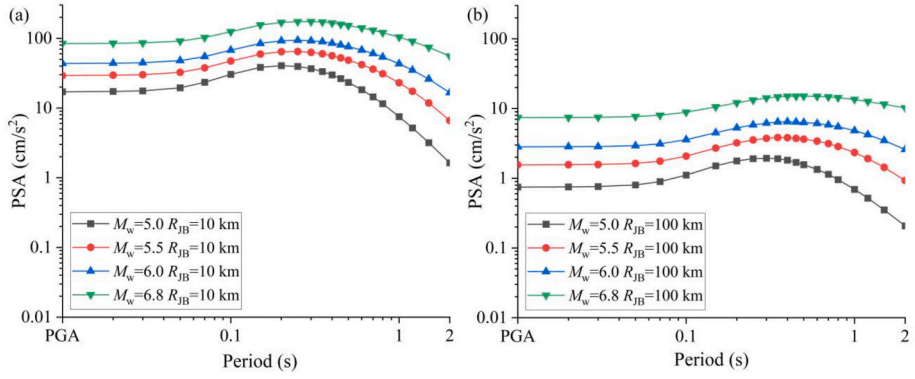


Fig. 6. Predicted response spectra of the GMM for various scenarios with (a) R_{JB} = 10 km and (b) R_{JB} = 100 km.

F_D = 8 km. The effect of anelastic attenuation shown in Fig. 6 is consistent with the results presented in Fig. 5. In addition, Fig. 6 demonstrates the relationship between earthquake magnitude and the predominant period (T_p) of the spectral response. T_p is observed to shift toward longer periods for larger-magnitude events. For example, T_p ranges from approximately 0.2 s for M_w 5.0 to 0.3 s for M_w 6.8. The phenomenon is observed in similar studies (e.g., Ref. [31]). As discussed by Kotha et al. [6], this phenomenon can be partially explained by the decreasing corner frequency with the increasing magnitude of the Brune [56] source spectrum model.

Compared with GMMs developed using recorded ground motion data from seismic events, the uncertainties in the simulation-based GMM might be lower where recorded data is sparse, such as very large earthquakes, rare site configurations, or close distances. Although simulated datasets can complement observed data and reduce epistemic uncertainty in poorly constrained regions, the simulation framework introduces its own uncertainties, primarily due to modeling assumptions. Additionally, non-ergodic models could be obtained through simulations for these regions compared to global models. Therefore, the uncertainties in simulated GMMs depend on the quality of the simulation framework and their validation against observed data.

Fig. 7 shows relatively small total standard deviation (σ) values, ranging from 0.30 to 0.41. The distribution of the inter-event (τ), intra-event (ϕ), and total (σ) uncertainties for PGA, PGV, and spectral amplitudes at various periods is compared with the nonparametric GMM proposed in Karimzadeh et al. [18]. The nonparametric GMM is developed using the XGBoost algorithm with the same simulation-based ground motion dataset. Across all spectral periods, the total uncertainties (σ) of the GMM developed in this study are larger than the results in Karimzadeh et al. [18], especially at short periods. This is primarily attributed to the notably low inter-event uncertainty, approximately 0.05, observed in the ML-based GMM developed by Karimzadeh et al. [18] across all period ranges. According to the comparison results in Fig. 7, the XGBoost algorithm tends to result in smaller uncertainties in terms of inter-event residuals compared to traditional mixed-effects regression methods. This is due to XGBoost's strong ability to capture global patterns in the data. The model benefits from regularization techniques, pruning, and iterative refinement, which help it generalize well and reduce overfitting, resulting in lower inter-event residuals. Compared with the empirical regression method, the low inter-event residual variation may result from using similar source mechanisms for the same region in the simulation. One of the key

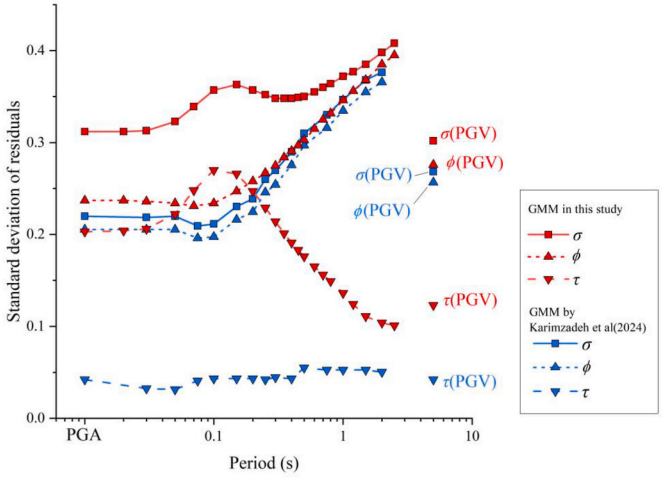


Fig. 7. Distribution of the inter-event (τ), intra-event (ϕ), and total (σ) uncertainties for PGA, PGV, and PSA ranging from 0.02 to 2.0 s. For comparison, variance estimates are compared between the nonparametric GMM by Karimzadeh et al. [18], which are developed using the same dataset with a machine learning-based method.

advantages of XGBoost is its ability to describe complex, nonlinear relationships, which traditional regression methods may fail to capture. This helps reduce the uncertainty of inter-events' residuals by better fitting the global structure of the data. The uncertainty of intra-event residual of the method is not reduced as inter-event. Intra-event variability is primarily influenced by path and site effects. Although the site amplification is removed from the simulation, the complex path effect will still impact the results. Due to the strong heterogeneity of the crustal medium, the microscale variations in the crust may not be adequately captured during simulations, either because of insufficient data or limited data resolution, as crustal models typically have a resolution on the kilometer scale. Additionally, when developing models, even ML approaches that rely on a single distance parameter may be overly simplistic and fail to capture the complex effects of wave propagation through the medium. In summary, XGBoost improves predictions at the event level by capturing global patterns more effectively. However, within-event residuals may remain similar between our proposed GMM, depending on the complexity of the input dataset and the model's tuning. Finally, it should be noted that in our current simulation framework, uncertainties are represented through the probability distributions assigned to source and path parameters and are captured via Monte Carlo simulations. However, epistemic uncertainties, such as those arising from model form assumptions and simplifications, are more difficult to isolate and are not explicitly incorporated. The misfit between simulated and observed ground motions reflects both modeling limitations and natural variability, but the scarcity of data makes it impractical to define a distinct modeling error term.

As shown in Fig. 7, there is a bump (i.e., increase) in the standard deviations centered near 0.1 s, which is controlled by the τ component. The bump is also observed in the results of previous studies, like in Kotha et al. [6] and Boore et al. [31]. In Boore et al. [31], some physical justifications are proposed for the bump: (1) variations in the source stress parameter for small magnitude earthquakes (effect absent for larger earthquakes) and (2) variations in zero distance kappa factor (κ0) for larger magnitude earthquakes (smaller magnitude events were not investigated). Further investigation of the short-period peak in τ is needed to verify and refine these hypotheses, which can be effectively achieved using simulation-based ground motion datasets.

6. Model performance evaluation and unbiasedness test

In this study, the performance of the developed GMM is assessed through four statistical indices, including *RMSE*, R^2, *r*, and *MAPE*. These metrics are computed as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\ln(y_i) - \ln(\hat{y}_i)\right)^2} \tag{12}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(\ln(y_i) - \ln(\hat{y}_i)\right)^2}{\sum_{i=1}^{n}\left(\ln(y_i) - \ln(\overline{y}_i)\right)^2} \tag{13}$$

$$r = \frac{\sum_{i=1}^{n}\left(\ln(y_i) - \ln(\overline{y})\right)\left(\ln(\hat{y}_i) - \ln(\overline{\hat{y}})\right)}{\sqrt{\sum_{i=1}^{n}\left(\ln(y_i) - \ln(\overline{y})\right)^2 \sum_{i=1}^{n}\left(\ln(\hat{y}_i) - \ln(\overline{\hat{y}})\right)^2}} \tag{14}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\ln(\hat{y}_i) - \ln(y_i)}{\ln(y_i)}\right| \tag{15}$$

where n is the number of samples; y denotes the target value of IM (herein simulated values); $\hat{y}$ signifies the predicted value; $\overline{y}$ and $\overline{\hat{y}}$ are the arithmetic means of the target and predicted values, respectively. To assess the accuracy of the proposed GMM, the estimated IMs in terms of

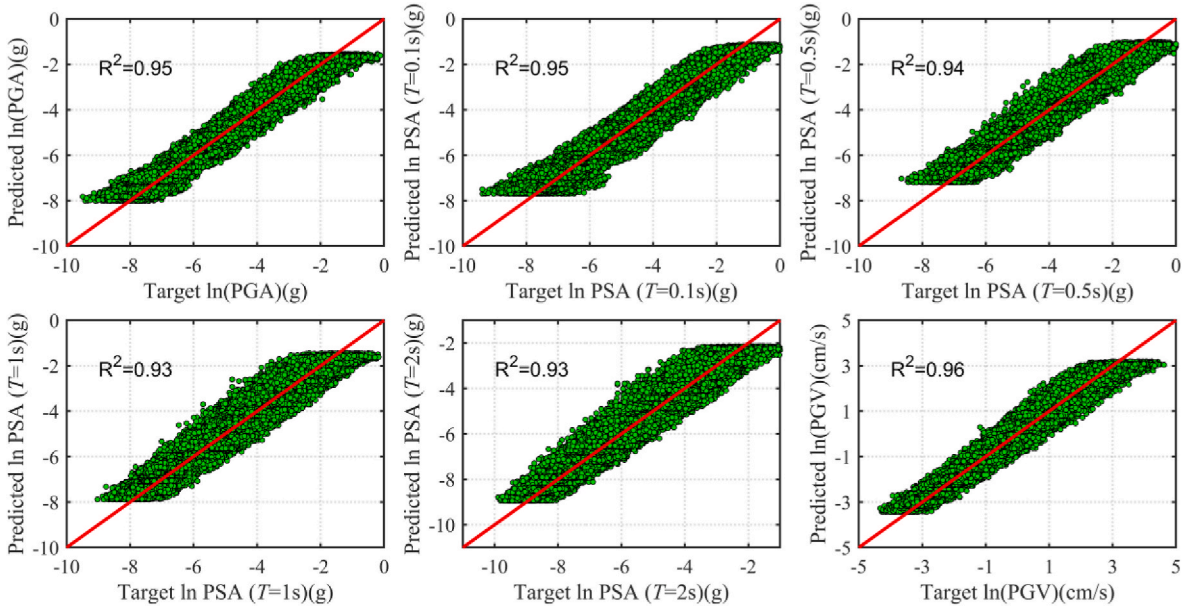


Fig. 8. Comparison of target (herein simulated) and predicted values form the developed GMM for all IMs. The coefficient of determination (R^2) is marked in the panels. The perfect fit line is illustrated with red lines. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

PGA, PGV, and PSA at different periods ranging from 0.02 to 2.0 s are compared with the corresponding target values. Among these metrics, R^2 quantifies the proportion of variance in the response variable that can be explained by the predictor variables. Pearson correlation coefficient r assesses the degree of linear relationship between the target and predicted IMs. The *RMSE* represents the average deviation between target and predicted values. A higher R^2 and r, coupled with a lower *RMSE*, indicate improved GMM performance.

Fig. 8 illustrates the model's effectiveness by comparing predicted and target values of various IMs on a natural logarithmic scale. The plots are presented with comparisons made against the ideal 1:1 fit line. A closer alignment with the ideal fit implies better model performance. The analysis demonstrates strong model efficacy for both datasets, as evidenced by an R^2 ranging from 0.93 to 0.96, indicating the model's high prediction accuracy.

Fig. 9 illustrates the statistical indices of the developed model, including *RMSE*, R^2, r, and *MAPE*. The results show that the model performs well across all IMs, with both r and R^2 values exceeding 0.92, indicating good model performance. The MAPE for PGA, PGV, and PSA is within 5 %~8 %. It is worth noting that when PGV is close to 1 cm/s, *MAPE* is unreliable due to the denominator approaching zero according to its definition in Eq. (15). To address this issue, MAPE for PGV is calculated under m/s unit. A slight decline in these two metrics is observed for PSA at longer periods, compared to PGA and PSA at shorter periods (less than 0.07 s), with R^2 values decreasing from 0.95 to 0.93 and r values decreasing from 0.98 to 0.97. The decrease in performance is further validated by the error term *RMSE*, which increases from 0.32 to 0.42. The decline in R^2 and r suggests that the developed model may perform better for PGA and spectral ordinates of shorter periods compared to PGV and spectral ordinates of longer periods, which should be taken into consideration when applying the model in practice. In studies of existing empirical GMMs, the performance decrease trend is more significant for PGV and PSA at longer period spectral ordinates [54]. As stressed by Motazedian and Atkinson [26], stochastic methodologies have struggled to accurately capture the coherent, long-period pulses that influence the characteristics of ground motions near faults, especially with respect to PGV and PSA at longer periods. Therefore, the applicable period range of our model is up to 2.0 s. The authors have also addressed this limitation in their recently published paper on the simulation of earthquake ruptures in the North Tabriz Fault in

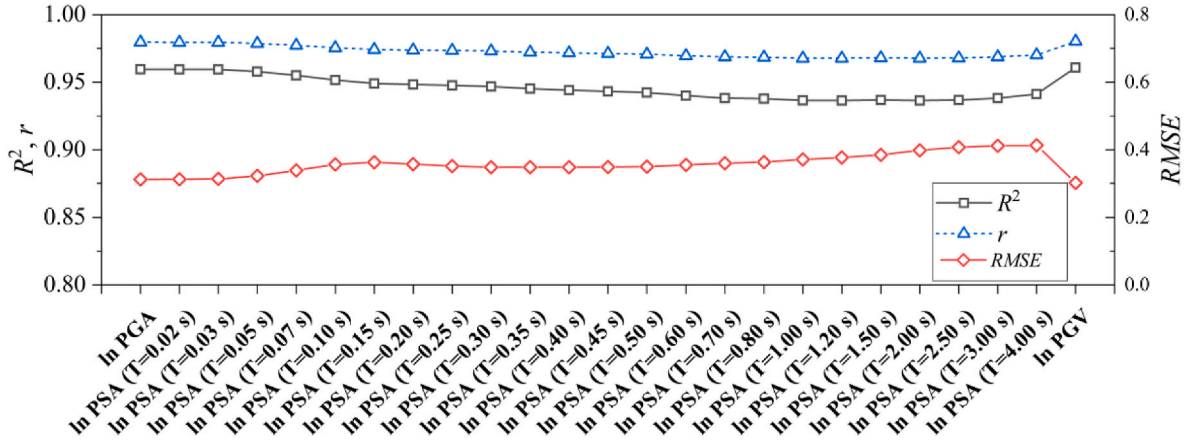


Fig. 9. Model performance metrics in terms of root-mean-squared-error (*RMSE*), coefficient of determination (R^2), Pearson correlation coefficient (r), and mean absolute percentage error (*MAPE*) (%).

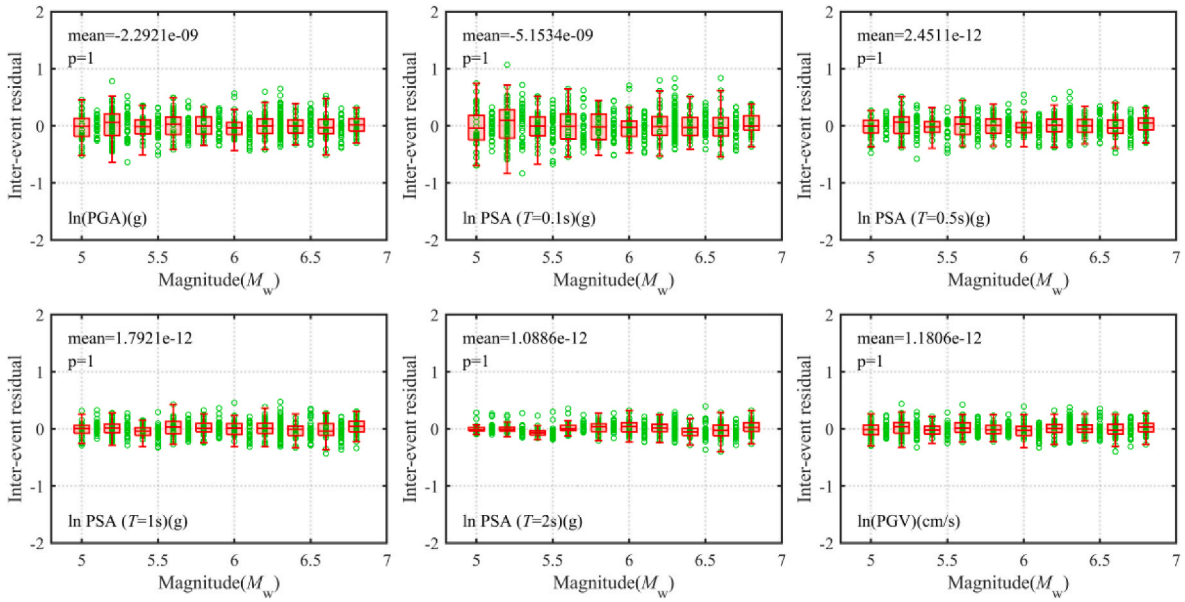


Fig. 10. Distribution of the inter-event residuals relative to magnitude, M_w, for the selected IMs.

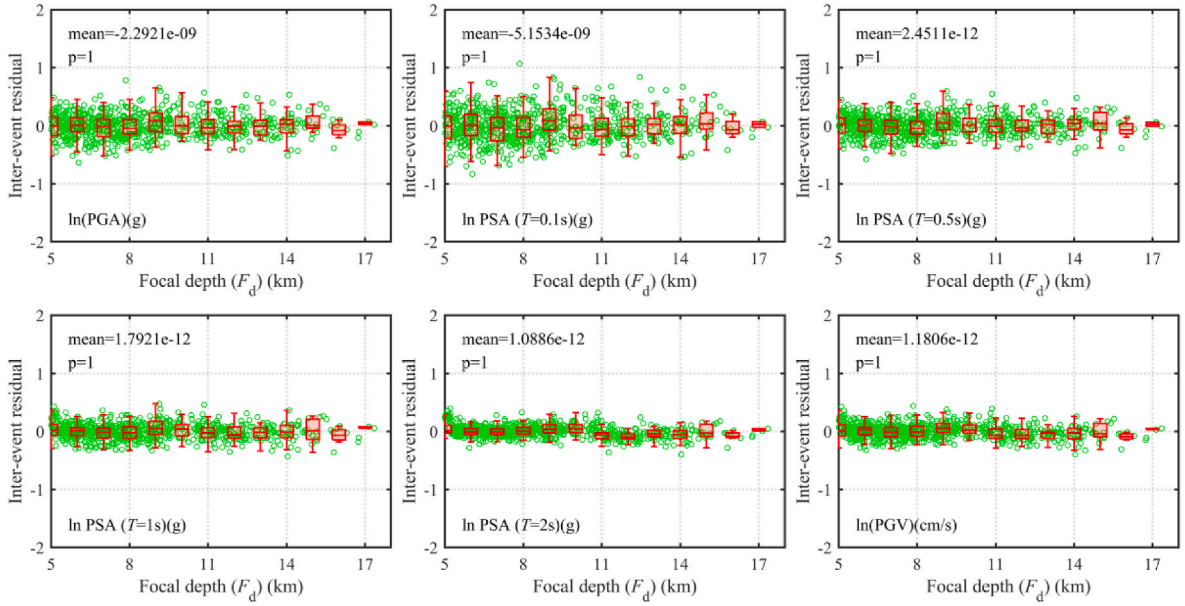


Fig. 11. Distribution of the inter-event residuals relative to focal depth, F_D, for the selected IMs.

Northwest Iran (Temiz et al. [57]), discussing the weakness of stochastic simulations in capturing long periods.

The total residuals range from 0.2 to 0.4, with higher values observed at longer spectral periods of PSA, which also concurs with previous studies. The model also demonstrates an acceptable range of uncertainty and performs well in comparison to existing models (e.g., NGA-West2 GMMs and Kotha et al. [6,54]). The potential bias of the model is evaluated by examining the inter-event and intra-event residuals in relation to source- and site-related parameters, M_w, F_D, and R_{JB}, respectively, as shown in Figs. 10–12. The median value and the 0.25/0.75 quartiles of the residuals regarding different bins of M_w, F_D, and R_{JB} are illustrated using the boxplot. No significant trends are observed in binned means across all considered IMs, indicating that the residuals are unbiased for all M_w, F_D, and R_{JB} ranges. The inter-event and intra-event residuals are consistent with previous research, ranging from −0.8 to 0.8 and −2.0 to 2.0, respectively. P-values are calculated at a significance level of 0.05 using a one-sample T-test, which is utilized to verify the null hypothesis of unbiased estimates. A p-value close to 1.0 for all considered IMs supports the assumption that the residual values come from a normal distribution with a mean value equal to zero, indicating that the GMM is independent of source-related or site-related explanatory variables for all frequency bands. Furthermore, unlike GMMs based on observation recordings, the uncertainty of residuals remains constant as magnitude increases or distance decreases, indicating a significant advantage of ground motion simulations over real datasets.

7. Comparison with existing ground motion models

The proposed GMM is further evaluated to assess its capability to capture physics-based characteristics of real earthquake behavior. To this end, results are compared for various magnitude–distance

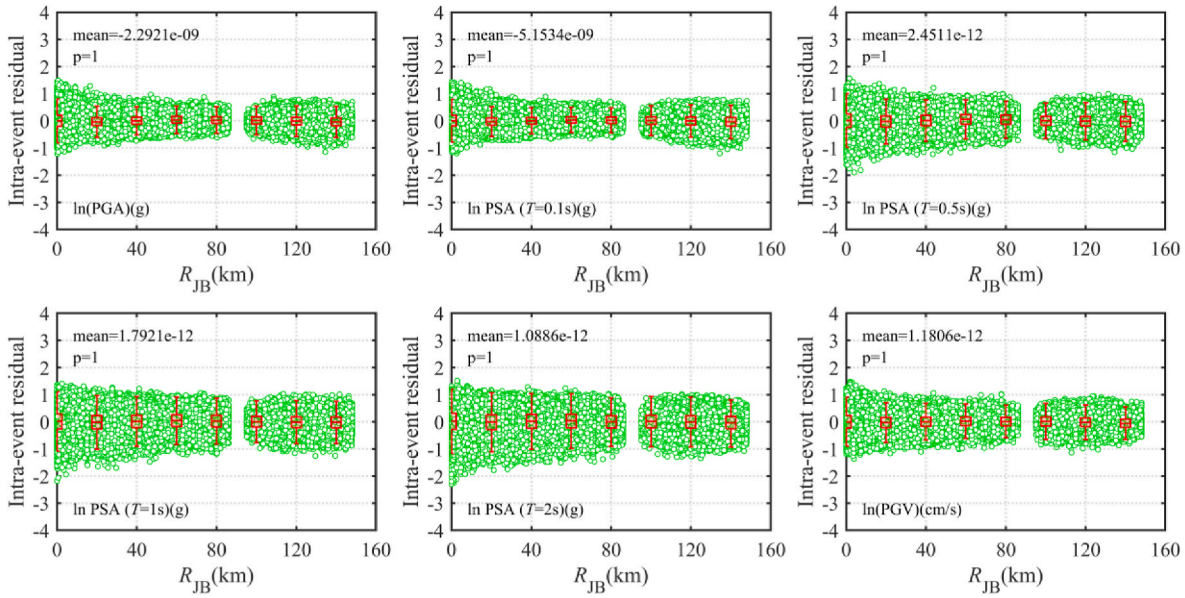


Fig. 12. Distribution of the intra-event residuals relative to distance, R_{JB}, for the selected IMs.

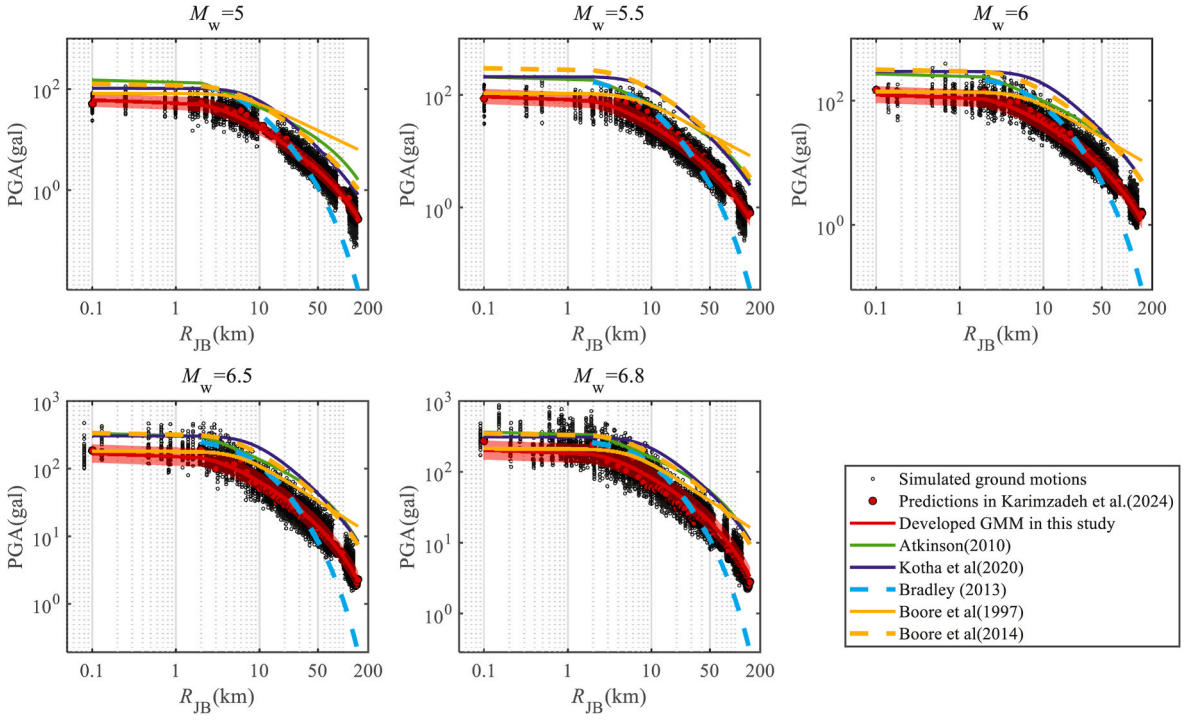


Fig. 13. Comparison of PGA against the nonparametric GMM of Karimzadeh et al. [18] and empirical GMMs of Atkinson [28], Bradley [29], Boore et al. [30], Boore et al. [31], and Kotha et al. [6].

combinations with a fixed F_D of 8.0 km. Figs. 13-16 presents the estimated PGA, PGV, and PSA at periods $T = 0.3$ s and 1.5 s for magnitudes ranging from M_w 5.0 to 6.8. This analysis spans a range of R_{JB} values. The findings reveal that as magnitude increases and distance decreases, there is a corresponding rise in PGA, PGV, and PSA across all period ranges. This indicates that the GMM effectively captures distance-dependent attenuation patterns.

Figs. 13-16 also compares the proposed GMM and several established GMMs from the literature, including those by Atkinson [28], Bradley [29], Boore et al. [30], Kotha et al. [6] and Boore et al. [31]. The GMMs from Atkinson [28] and Bradley et al. [29] are selected due to their compatibility with the distinct tectonic characteristics of the studied Azores region. The GMM by Atkinson [28] was specifically developed for volcanic earthquakes in Hawaii, which is explicitly tailored for volcanic regions. The GMM proposed by Bradley et al. [29] considers the complex anelastic attenuation in the Taupo Volcanic Zone of New Zealand. Although these two GMMs are not global models, they are developed for volcanic island settings and share comparable seismic and tectonic characteristics with the Azores. The GMM proposed by Boore et al. [30] is included due to its attenuation characteristics closely matching that observed in the Azores region, according to the study by Oliveira et al. [58]. The GMM in Kotha et al. [6] is included as it is

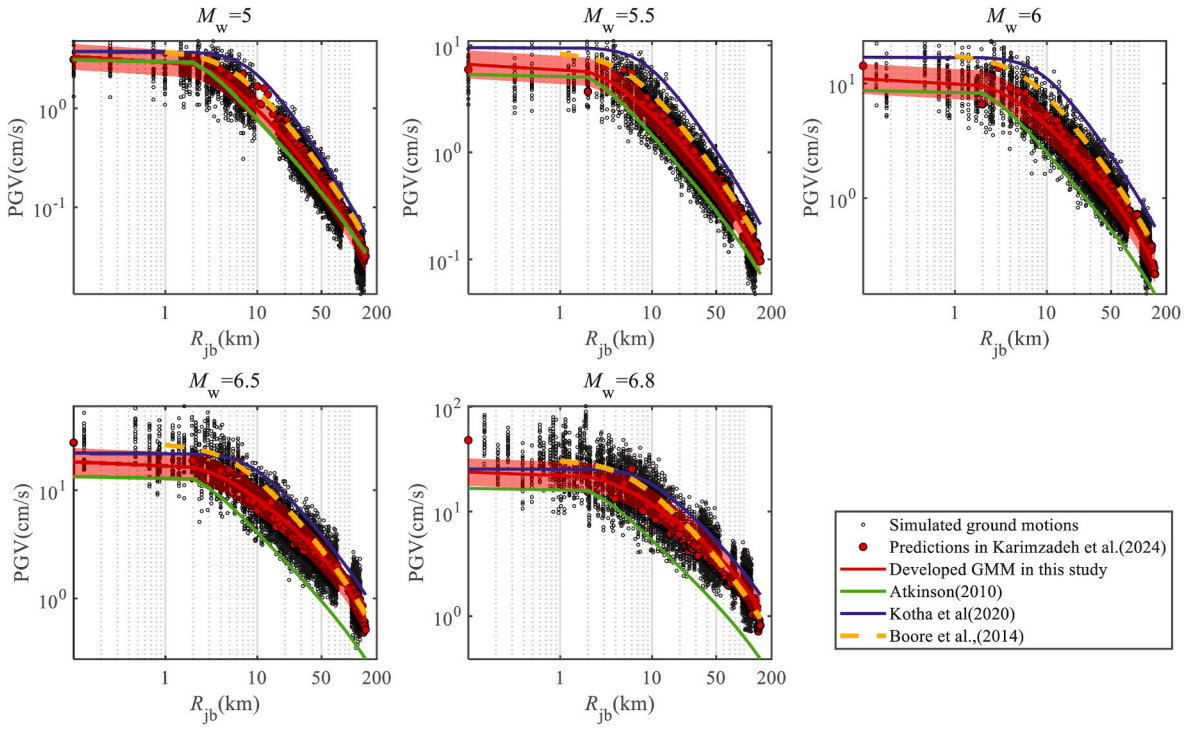


Fig. 14. Comparison of PGV against the nonparametric GMM of Karimzadeh et al. [18] and empirical GMMs of Atkinson [28], Kotha et al. [6], and Boore et al. [31].

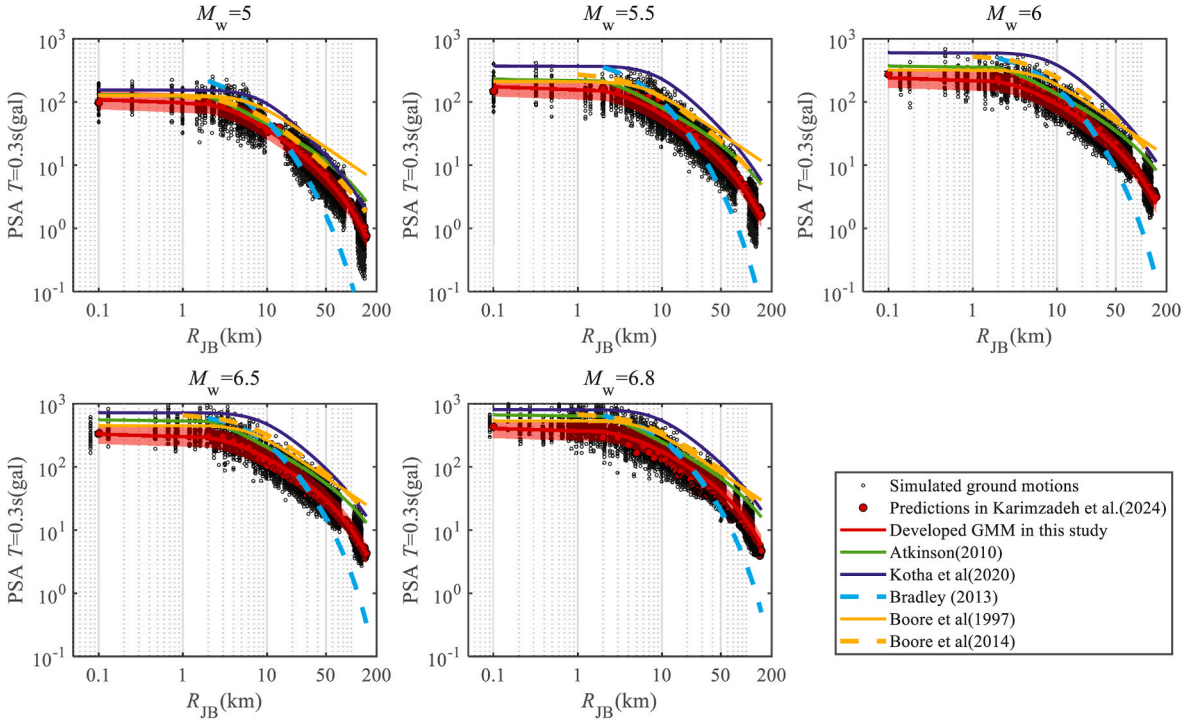


Fig. 15. Comparison of PSA ($T = 0.3$s) against the nonparametric GMM of Karimzadeh et al. [18] and empirical GMMs of Atkinson [28], Bradley [29], Boore et al. [30], Boore et al. [31] and Kotha et al. [6].

developed for the active shallow crust in ESHM20 (Danciu et al. [59, 60]). Lastly, our model is compared with one global NGA-West2 GMM for shallow crustal seismicity, specifically the Boore et al. [31] model.

Additionally, the GMM is evaluated against the region-specific nonparametric model introduced by Karimzadeh et al. [18]. This comparative analysis aims to place the performance of the proposed model within the broader context of existing modeling approaches and provides insights into its potential across various seismic settings. The comparison reveals discrepancies between the proposed region-specific model and global models, suggesting the influence of distinct regional seismological characteristics, as also highlighted by Karimzadeh et al. [18]. This underlines the importance of region-specific data in GMM

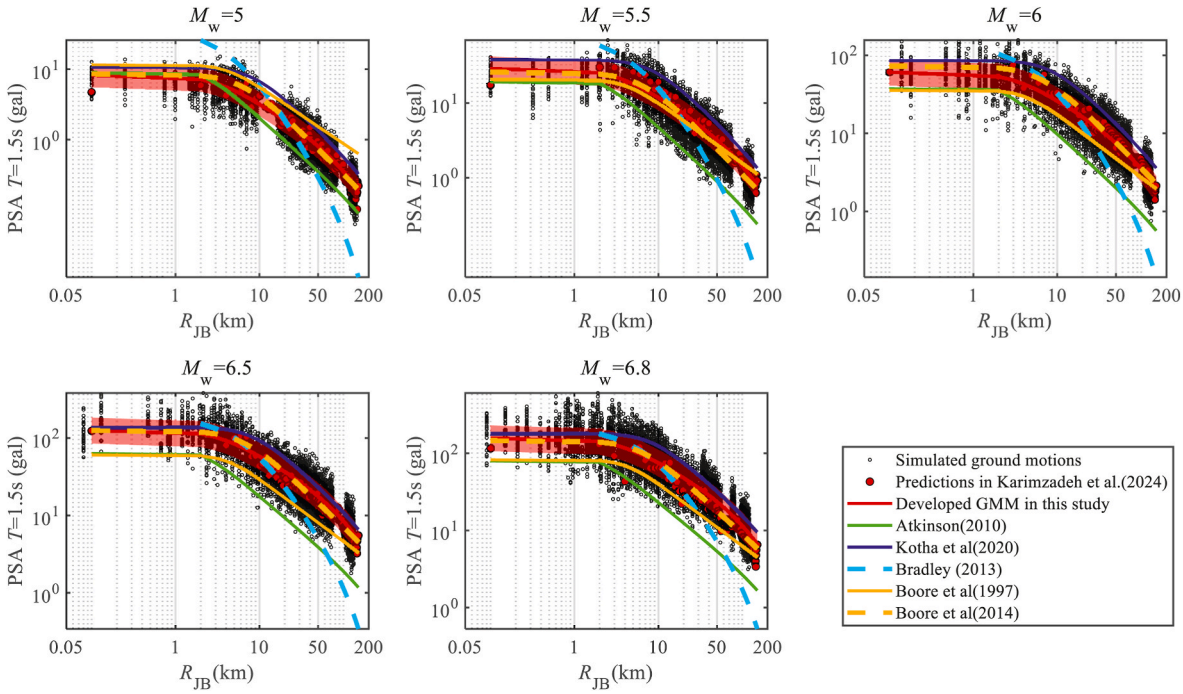


Fig. 16. Comparison of PSA (T = 1.5 s) against the nonparametric GMM of Karimzadeh et al. [18] and empirical GMMs of Atkinson [28], Bradley [29], Boore et al. [30], Boore et al. [31] and Kotha et al. [6].

studies for this area. Conversely, the consistency of the proposed model with the local nonparametric model indicates the model's efficacy in capturing the region's nonlinear ground motion behavior. This evidence supports the suitability of the proposed empirical formula for representing ground motion phenomena specific to the study region.

Comparative analysis with established empirical GMMs shows that simulated attenuation in the Azores region, especially for PGA, decays faster than models such as Boore et al. [30], which is consistent with previous findings. The predicted results by Kotha et al. [6] are larger overall than our model for all the IMs considered, while the decay tendency is similar to our model regarding different magnitude bins. The results in Bradley [29] show a significantly faster decay rate than all the other empirical models. For PGA and PSA(T = 0.3s) at short distances (R_{JB} < 5 km), the developed model has smaller amplitudes than results using GMMs in Ref. [28,30,31,6]. In terms of PGV at short distances (R_{JB} < 5 km), the developed model has smaller amplitudes than Kotha et al. [6].

The amplitude difference is smaller when the magnitude is larger than M_w 6.1. For a magnitude less than M_w5.6, the developed model has similar PGV amplitudes compared to the model in Atkinson [28]. In other magnitude cases, the developed model has larger amplitudes than Atkinson [28]. For small magnitude events (M_w 5.0), the PGV prediction results are generally consistent with the Atkinson [28] model but slightly lower than those of Boore et al. [31] and Kotha et al. [6] models. This underestimation becomes more pronounced for magnitudes M_w 5.5 and 6.0. A similar trend is observed for PSA at long period T = 1.5 s.

The differences in amplitude and attenuation rate compared to other GMMs are primarily attributed to the complex seismic wave propagation and heterogeneities in the volcanic environment of the studied Azores islands. As shown by Molkenthin et al. [61] and Douglas and Jousset [62], in a stochastic simulation framework, PSA for small to moderate events at high oscillator frequencies of 10 Hz is primarily controlled by kappa. High kappa values correspond to low-frequency energy, resulting in lower ground motion for the same magnitude and epicentral distance. The relatively high kappa value for the Azores islands, 0.075 ± 0.02s (Table 2), is consistent with values measured in seismically active regions like California in the western United States, having kappa around 0.04–0.07 s [63]. In Europe and the Mediterranean, the average kappa value is estimated as 0.0308 s and approximately 0.0457 and 0.0261 s for Türkiye and Italy, respectively, according to analysis of the pan-European strong motion database by Bora et al. [64]. The kappa value for the Azores islands is higher than that in mainland Europe. The GMM in Boore et al. [30] was regressed using the earthquake data from the Western United States with high kappa. This partially explains why the PGA amplitudes predicted by the proposed GMM are lower than those from the model in Kotha et al. [6], and closer to the results of Boore et al. [30]. The higher kappa value for the Azores islands could also partially explain why the predicted PGA amplitude using our model is lower than the global GMM results proposed in Boore et al. [31]. Since kappa primarily controls the high oscillator frequencies, the differences in predicted amplitudes for long-period PSA, like PSA (T = 1.5 s), are insignificant across GMMs.

Additionally, the stochastic simulations for the Azores islands consider only the normal faulting events, which represent the predominant faulting mechanism in the islands. In the development of New Zealand GMMs by Bradley [29], significant over-predictions of normal faulting events at short periods were observed for global GMMs (e.g., the CB08 model [65] in the NGA-West1 project). This is primarily due to the limited proportion of recordings from normal events in the NGA dataset. According to Kotha et al. [6], only non-subduction events were retained in GMM regression, and the normal faulting events across the Azores islands were excluded. Moreover, the style-of-faulting random effect was not incorporated in the regression of the Kotha et al. [6] model because it was argued that style-of-faulting factors cannot be quantified using isotropic factors without azimuthal variations [66]. This may partially explain the notable over-prediction of normal faulting events in the Azores islands for Kotha et al. [6] GMM at short periods.

The quality factor (Q) used for finite fault simulation of ground motions in the Azores islands is set as $Q(f) = (76 \pm 11)\, f^{0.69\pm0.09}$ (Table 2). The low Q_o value, 76 ± 11, agrees with those from other volcanic areas, like Mount Cameroon volcanic region in West Africa (Q_o = 65, Ref [67]), Long Valley Caldera region in California (Mayeda et al. [68]) and Nevado del Ruiz Volcano in Colombia (Q_o = 30, Londono [69]). The relatively low Q_o explains why the decay rate is more rapid

than other non-volcanic GMMs, including Boore et al. [30], Boore et al. [31] and Kotha et al. [6]. This also explains why the apparent anelastic attenuation term in our proposed model (a_6) is significantly lower than the corresponding term in global NGA-West2 GMM by Boore et al. [31], as shown in Fig. 5(c). For the Taupo Volcanic Zone in New Zealand, the value of Q_o is estimated at 36.0 (Bradley [29]). The Q_o value is even lower than that in the Azores islands, indicating stronger anelastic attenuation than in the Azores. This corresponds to the steeper decay rate observed in the Bradley13 GMM. The PGV and PSA (T = 1.5 s) decay rate with distance is not as significantly different from that of non-volcanic region GMMs (such as Kotha et al. [6]) as it is for PGA and short-period PSA. This suggests that the difference in attenuation slopes due to Q_0 variations is more pronounced in the high-frequency range and becomes less significant at moderate to long periods.

In Karimzadeh et al. [18], simulations were compared against observed acceleration time histories, PSA, PGA, and PGV at four stations in the 1998 Faial earthquake, which are the only instrumental recordings of significant events in our study region. The comparison yielded Goodness-of-Fit (GOF) scores ranging from 77 to 84 (as shown in Karimzadeh et al. [18]), based on the metric developed by Olsen and Mayhew [70], confirming that the input parameters, when calibrated with regional site amplification factors, provide reliable performance. It is important to emphasize that our proposed GMM is developed for bedrock conditions, and as such, site amplification effects were not incorporated in generating the simulation-based dataset. Therefore, a direct comparison with observed records from the 1998 Faial earthquake is not appropriate, as none of the four recording stations were situated on rock or very stiff soil. Compared to these observations, the bedrock-only GMM is expected to systematically underestimate both PGA and PSA values, due to the absence of site amplification in the model.

Furthermore, the region exhibits unique site-specific effects driven by its distinctive geotechnical and geological features, including volcanic formations [71]. As such, reliable ground motion predictions require site amplification models that are calibrated through detailed, site-specific response analyses, accounting for local soil profiles. In ongoing work, we are developing a regionally calibrated site amplification model tailored to the area's complex subsurface conditions. This effort will enable more accurate incorporation of the site effects in future simulation-based GMMs for scenario earthquakes across all relevant faults, thereby extending the analysis beyond the limitations of a single-event simulation.

8. Conclusions

This study presents the development of a GMM for the Azores Plateau in Portugal utilizing a simulated, homogeneous dataset specific to the region. The dataset was generated by Karimzadeh et al. [18] using a stochastic finite-fault approach that incorporates uncertainties in fault ruptures and path attenuation across various earthquake scenarios. Subsequently, a mixed-effects algorithm is employed to construct a parametric GMM estimating IMs, including peak ground acceleration (PGA), peak ground velocity (PGV), and pseudo-spectral acceleration (PSA) with a 5 % damping ratio over a period range of 0.02–2.0 s. The model uses key input parameters such as moment magnitude (M_w), Joyner-Boore distance (R_{JB}), and focal depth (F_D).

The study develops GMM using a nonlinear regression algorithm known for its high accuracy, adaptability, and computational speed. The model's robustness is assessed by analyzing potential bias in the residuals through examination of inter-event and intra-event variability linked to source and site parameters. Results show that the residuals are statistically unbiased across all considered IMs, magnitude, and distance ranges. The proposed GMM achieves high agreement between simulated and predicted values for various IMs, such as PGA, PGV, and PSA, at different periods. The developed model shows an acceptable range of uncertainty and performs favorably to existing models, as indicated by the inter-event, intra-event, and total residuals for all IMs. The findings reveal that the variations in earthquake magnitude and distance influence spectral ordinates in a way consistent with established earthquake physical expectations, supporting the model's validity for application in the Azores region.

The model is compared with several existing GMMs, including those developed by Boore et al. [30], Atkinson [28], Bradley [29], Kotha et al. [6] and global model in Boore et al. [31]. Moreover, the GMM is assessed against the region-specific ML-based model proposed by Karimzadeh et al. [18]. This evaluation covers different IMs such as PGA, PGV, and PSA at periods T = 0.3 s and T = 1.5 s. The results highlight notable differences between the proposed model and global models, pointing to the influence of the Azores' distinct seismological characteristics, as also emphasized by Karimzadeh et al. [41]. This finding reinforces the need for region-specific GMM in seismic hazard assessment for this region. Comparative analysis with established empirical GMMs shows that the simulated attenuation within the Azores region, especially for PGA, decays faster compared to models such as Boore et al. [30,31], which is consistent with previous findings. On the other hand, the amplitude of predicted results by Kotha et al. [6] is larger than our model overall for all considered IMs, although its decay trend with respect to magnitude bins is similar. These differences are consistent with recent studies [72–78], which emphasize significant variations in ground-motion behavior across tectonic environments, earthquake types, and motion components. The strong agreement between the proposed model and the local ML-based nonparametric model further validates its ability to capture the nonlinear characteristics of ground motion in the Azores region.

The stochastic simulation method enhances seismic hazard assessments by providing a non-ergodic, region-specific model for areas with limited recorded motions and insufficient detailed information (e.g., velocity structure, source model) to perform physics-based simulations. The approach offers a more practical approach by capturing physical properties through simplified source, path, and site parameters. Because the input model parameters for the stochastic finite-fault simulation approach are entirely derived from real dataset behavior [42], the developed model represents the real data more accurately than global models. Moreover, Monte Carlo techniques applied in developing ground motion datasets and GMMs enable robust probabilistic modeling by repeatedly sampling input parameters. Moving forward, future research should integrate physics-based simulations with stochastic approaches to improve the representation of low-frequency ground motion and enhance the model's overall precision and applicability.

When defining site classes according to Eurocode 8 [79], our proposed GMM applies only to firm-rock conditions (Site Class A or $V_{S30} \geq$ 800 m/s), limiting its direct applicability to site-specific seismic hazard analyses. In line with standard engineering practice and seismic design codes, we propose adjusting the GMM output using amplification factors provided in relevant design codes. For instance, Eurocode 8 [79] offers spectral shape and amplification factors for various site classes, which can be applied to the bedrock GMM to estimate ground motions at the soil site. This approach provides a practical and code-compliant means of incorporating site effects in regions where site-specific data is unavailable. For users aiming to account for site amplification across a range of V_{S30} values, we recommend adjusting our bedrock median spectra using established empirical site amplification models, such as the Seyhan & Stewart [80] model adopted in the NGA-West2 GMM. However, it is important to note that this method is inherently approximate, as it reflects the generalized behavior of global datasets rather than local conditions. In particular, the studied volcanic region presents distinct site effects due to its unique geotechnical characteristics, which may not be adequately captured by globally derived models. If feasible, results should be adjusted using site amplification models calibrated through site-specific soil response analyses that account for local soil profiles and dynamic soil properties. This represents a future step in our efforts to develop region-specific site amplification factors.

Furthermore, for the volcanic region under study, we are actively contributing to the update of the national annex, intending to refine the site amplification factors based on region-specific geological and geophysical data. This ongoing work will enhance the reliability of site-adjusted seismic hazard assessments in future studies.

Finally, in this study, uncertainties were addressed through probabilistic distributions assigned to source and path parameters, propagated via Monte Carlo simulations. This approach effectively captures aleatory variability; however, epistemic uncertainties, particularly those arising from model assumptions and simplifications, remain unquantified due to limited observational data. To address these limitations, future research should explore the integration of alternative modeling frameworks and hybrid simulation strategies. In particular, advanced machine learning techniques, such as conditional variational autoencoders, offer promising tools for representing epistemic uncertainties in data-scarce regions and improving the robustness of ground motion predictions.

CRediT authorship contribution statement

Kun Ji: Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Methodology, Investigation, Formal analysis. **Shaghayegh Karimzadeh:** Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Saman Yaghmaei-Sabegh:** Writing – review & editing, Supervision, Resources, Investigation. **Ruibin Hou:** Writing – review & editing, Supervision, Methodology, Investigation. **Alexandra Carvalho:** Writing – review & editing. **Paulo B. Lourenço:** Writing – review & editing, Supervision, Resources, Funding acquisition.

Data availability

The GMM code is available at https://github.com/JIKUN1990/GMMs. Users can input parameters for scenario earthquakes (moment magnitude, M_w, Joyner-Boore distance, R_{JB}, and focal depth, F_D) to generate IMs, including peak ground acceleration (PGA), peak ground velocity (PGV), and pseudo-spectral acceleration (PSA) at periods between 0.02 and 2.0 s. The simulated earthquake catalog and IMs, including PGA, PGV, and PSA at periods between 0.02 and 2.0 s, are available at https://doi.org/10.5281/zenodo.14348200.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study has been partly funded by the STAND4HERITAGE project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant agreement No. 833123), as an Advanced Grant. This work has been partly financed by the Fundamental Research Funds for the Central Universities (B240201087). This work was partly financed by FCT/MCTES through national funds (PIDDAC) under the R&D Unit Institute for Sustainability and Innovation in Structural Engineering (ISISE), under reference UIDB/04029/2020 (doi.org/10.54499/UIDB/04029/2020), and under the Associate Laboratory Advanced Production and Intelligent Systems ARISE under reference LA/P/0112/2020. Finally, this work has been partly financed by the Chinese National Natural Science Fund (52378506).

Data availability

It is already declared in the manuscript. The data is available in Zenodo, and the model is available in Github.

References

[1] Martino P, Daniele E, Thomas K, Alice S, Aneta F, Christina C. Atlas of the human planet 2017: global exposure to natural hazards. EUR 28556 EN 2017. https://doi.org/10.2760/19837.
[2] Baker JW, Bradley B, Stafford P. Seismic hazard and risk analysis. Cambridge University Press; 2021.
[3] Bozorgnia Y, Abrahamson NA, Atik LA, Ancheta TD, Atkinson GM, Baker JW, Baltay A, Boore DM, Campbell KW, Chiou BS, Darragh R. NGA-West2 research project. Earthq Spectra 2014;30(3):973–87. 2014 Aug.
[4] Mousavi M, Ansari A, Zafarani H, Azarbakht A. Selection of ground motion prediction models for seismic hazard analysis in the Zagros region, Iran. J Earthq Eng 2012;16(8):1184–207.
[5] Darzi A, Zolfaghari MR, Cauzzi C, Fäh D. An empirical ground-motion model for horizontal PGV, PGA, and 5% damped elastic response spectra (0.01–10 s) in Iran. Bull Seismol Soc Am 2019;109(3):1041–57.
[6] Kotha SR, Weatherill G, Bindi D, Cotton F. A regionally-adaptable ground-motion model for shallow crustal earthquakes in Europe. Bull Earthq Eng 2020;18(9):4091–125.
[7] Parker GA, Stewart JP, Boore DM, Atkinson GM, Hassani B. NGA-subduction global ground motion models with regional adjustment factors. Earthq Spectra 2022;38(1):456–93.
[8] Jung SJ, Yee E. Compatible ground motion models for South Korea using moderate earthquakes. Appl Sci 2024;14(3):1182.
[9] Lavrentiadis G, Abrahamson NA, Nicolas KM, Bozorgnia Y, Goulet CA, Babič A, Macedo J, Dolšek M, Gregor N, Kottke AR, Lacour M. Overview and introduction to development of non-ergodic earthquake ground-motion models. Bull Earthq Eng 2023;21(11):5121–50.
[10] Douglas J. Ground motion prediction equations 1964–2021. Department of Civil and Environmental Engineering, Imperial College London; 2021.
[11] Kubo H, Naoi M, Kano M. Recent advances in earthquake seismology using machine learning. Earth Planets Space 2024;76(1):36.
[12] Bindi D, Pacor F, Luzi L, Puglia R, Massa M, Ameri G, Paolucci R. Ground motion prediction equations derived from the Italian strong motion database. Bull Earthq Eng 2011;9:1899–920.
[13] Yazdani A, Kowsari M. Earthquake ground-motion prediction equations for northern Iran. Nat Hazards 2013;69:1877–94.
[14] Boore DM, Stewart JP, Skarlatoudis AA, Seyhan E, Margaris B, Theodoulidis N, Scordilis E, Kalogeras I, Klimis N, Melis NS. A ground-motion prediction model for shallow crustal earthquakes in Greece. Bull Seismol Soc Am 2021;111(2):857–74.
[15] Mohammadi A, Karimzadeh S, Banimahd SA, Ozsarac V, Lourenço PB. The potential of region-specific machine-learning-based ground motion models: application to Turkey. Soil Dynam Earthq Eng 2023;172:108008.
[16] Mohammadi AH, Hussaini SMS, Caicedo D, Karimzadeh S, Lourenço PB. Nonparametric ground motion models of cumulative absolute velocity and peak ground velocity for the Italian dataset. In: International conference on energy and environmental science. Cham: Springer Nature Switzerland; 2023. p. 43–54.
[17] Karimzadeh S, Mohammadi A, Hussaini SMS, Caicedo D, Lourenço PB. ANN-based ground motion model for Turkey using stochastic simulation of earthquakes. Geophys J Int 2024;236(1):413–29.
[18] Karimzadeh S, Mohammadi A, Salahuddin U, Carvalho A, Lourenço PB. Backbone ground motion model through simulated records and XGBoost machine learning algorithm: an application for the Azores plateau (Portugal). Earthq Eng Struct Dynam 2024;53(2):668–93.
[19] Hussaini SMS, Caicedo D, Mohammadi A, Karimzadeh S, Lourenço PB. Nonparametric ground motion models of arias intensity and significant duration for the Italian dataset. J Phys Conf 2024, June;2647(6):062001. IOP Publishing.
[20] Rezaeian S, Sun X. Stochastic ground motion simulation. In: Beer M, Kougioumtzoglou I, Patelli E, Au IK, editors. Encyclopedia of earthquake engineering. Berlin, Heidelberg: Springer; 2014. https://doi.org/10.1007/978-3-642-36197-5_239-1.
[21] Anbazhagan P, Kumar A, Sitharam TG. Ground motion prediction equation considering combined dataset of recorded and simulated ground motions. Soil Dynam Earthq Eng 2013;53:92–108.
[22] Yenier E, Atkinson GM. Regionally adjustable generic ground-motion prediction equation based on equivalent point-source simulations: application to central and eastern North America. Bull Seismol Soc Am 2015;105(4):1989–2009.
[23] Sandıkkaya MA, Akkar S, Kale Ö, Yenier E. A simulation-based regional ground-motion model for Western Turkiye. Bull Earthq Eng 2023;21(7):3221–49.
[24] Abrahamson NA, Youngs RR. A stable algorithm for regression analyses using the random effects model. Bull Seismol Soc Am 1992;82(1):505–10.
[25] Atkinson GM, Assatourians K. Implementation and validation of EXSIM (a stochastic finite-fault ground-motion simulation algorithm) on the SCEC broadband platform. Seismol Res Lett 2015;86(1):48–60.
[26] Motazedian D, Atkinson GM. Stochastic finite-fault modeling based on a dynamic corner frequency. Bull Seismol Soc Am 2005;95(3).
[27] Boore DM. Comparing stochastic point-source and finite-source ground-motion simulations: SMSIM and EXSIM. Bull Seismol Soc Am 2009;99(6):3202–16.
[28] Atkinson GM. Ground-motion prediction equations for Hawaii from a referenced empirical approach. Bull Seismol Soc Am 2010;100:751–61.
[29] Bradley BA. A New Zealand-specific pseudospectral acceleration ground-motion prediction equation for active shallow crustal earthquakes based on foreign models. Bull Seismol Soc Am 2013;103:1801–22.

[30] Boore DM, Joyner WB, Fumal TE. Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes: a summary of recent work. Seismol Res Lett 1997;68(1):128–53.
[31] Boore DM, Stewart JP, Seyhan E, Atkinson GM. NGA-West2 equations for predicting PGA, PGV, and 5% damped PSA for shallow crustal earthquakes. Earthq Spectra 2014;30(3):1057–85.
[32] Hipólito A, Madeira J, Carmo R, Gaspar JL. Neotectonics of Graciosa Island (Azores): a contribution to seismic hazard assessment of a volcanic area in a complex geodynamic setting. Ann Geophys 2013;56(6):1–18.
[33] Carmo R, Madeiram J, Ferreira T, Queiroz G, Hipólito A. Chapter 6 volcano-tectonic structures of São Miguel island, Azores. Geol Soc Spec Publ 2015;44(1):65–86. London, Memoirs.
[34] Madeira J, Brum da Silveira A, Hipólito A, Carmo R. Chapter 3 active tectonics in the central and eastern Azores islands along the Eurasia–Nubia boundary: a review. Geol Soc Spec Publ 2015;44(1):15–32. London, Memoirs.
[35] Wells DL, Coppersmith KJ. New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. Bull Seismol Soc Am 1994;84(4):974–1002.
[36] Atkinson GM, Goda K, Assatourians K. Comparison of nonlinear structural responses for accelerograms simulated from the stochastic finite-fault approach versus the hybrid broadband approach. Bull Seismol Soc Am 2011;101(6):2967–80.
[37] Assatourians K, Atkinson GM. EXSIM12: a stochastic finite-fault computer program in FORTRAN. Available at: http://www.seismotoolbox.ca, ; 2012. November 2024.
[38] Beresnev IA, Atkinson GM. FINSIM-A FORTRAN program for simulating stochastic acceleration time histories from finite faults. Seismol Res Lett 1998;69. https://doi.org/10.1785/gssrl.69.1.27.
[39] Boore DM. Comparing stochastic point-source and finite-source ground-motion simulations: SMSIM and EXSIM. Bull Seismol Soc Am 2009;99(6):3202–16.
[40] Karimzadeh S, Lourenço PB. Stochastic ground motion simulation of the 9th of july 1998 faial earthquake (Azores, North atlantic) authorea preprints. 2022.
[41] Karimzadeh S, Hussaini SMS, Funari MF, Lourenço PB. On the effect of different code-based ground motion selection approaches for the estimation of the seismic demand of masonry structures by using real ground motion data set. Authorea Preprints 2021. https://doi.org/10.1002/essoar.10509375.1.
[42] Carvalho A, Reis C, Vales D. Source and high-frequency decay parameters for the Azores region for stochastic finite-fault ground motion simulations. Bull Earthq Eng 2016;14:1885–902.
[43] Karimzadeh S, Funari MF, Szabó S, Hussaini SS, Rezaeian S, Lourenço PB. Stochastic simulation of earthquake ground motions for the seismic assessment of monumental masonry structures: source-based vs site-based approaches. Earthq Eng Struct Dynam 2024;53(1):303–30.
[44] Bernardo V, Karimzadeh S, Caicedo D, Hussaini SS, Lourenço PB. Fragility-based seismic assessment of traditional masonry buildings on Azores (Portugal) using simulated ground-motion records. Earthq Spectra 2024;40(4):2836–61.
[45] Jaimes MA, Lermo J, García-Soto AD. Ground-motion prediction model from local earthquakes of the Mexico basin at the hill zone of Mexico City. Bull Seismol Soc Am 2016;106(6):2532–44.
[46] Donahue JL, Abrahamson NA. Simulation-based hanging wall effects. Earthq Spectra 2014;30(3):1269–84.
[47] Weatherill G, Crowley H, Roullé A, Tourlière B, Lemoine A, Gracianne C, Kotha SR, Cotton F. Modelling site response at regional scale for the 2020 European Seismic Risk Model (ESRM20). Bull Earthq Eng 2023;21(2):665–714.
[48] Loviknes K, Cotton F, Weatherill G. Exploring inferred geomorphological sediment thickness as a new site proxy to predict ground-shaking amplification at regional scale: application to Europe and eastern Türkiye. Nat Hazards Earth Syst Sci 2024;24(4):1223–47.
[49] Darzi A, Halldorsson B, Cotton F, Rahpeyma S. Nationwide frequency-dependent seismic site amplification models for Iceland. Soil Dynam Earthq Eng 2024;183:108798.
[50] Huang YN, Whittaker AS, Luco N. NEHRP site amplification factors and the NGA relationships. Earthq Spectra 2010;26(2):583–93.
[51] Kawase H, Nagashima F, Nakano K, Mori Y. Direct evaluation of S-wave amplification factors from microtremor H/V ratios: double empirical corrections to "Nakamura" method. Soil Dynam Earthq Eng 2019;126:105067.
[52] Bazzurro P, Cornell CA. Ground-motion amplification in nonlinear soil sites with uncertain properties. Bull Seismol Soc Am 2004;94(6):2090–109.
[53] Campbell KW, Bozorgnia Y. NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra. Earthq Spectra 2014;30(3):1087–115.
[54] Kotha SR, Bindi D, Cotton F. Partially non-ergodic region specific GMPE for Europe and Middle-East. Bull Earthq Eng 2016;14:1245–63.
[55] Lagarias JC, Reeds JA, Wright MH, Wright PE. Convergence properties of the Nelder-Mead simplex method in low dimensions. SIAM J Optim 1998;9(1):112–47.
[56] Brune JN. Tectonic stress and the spectra of seismic shear waves from earthquakes. J Geophys Res 1970;75(26):4997–5009.
[57] Temiz C, Hussaini SMS, Karimzadeh S, Askan A, Lourenço PB. Seismic scenario simulation and ANN-based ground motion model development on the North Tabriz Fault in Northwest Iran. J Seismol 2024. https://doi.org/10.1007/s10950-024-10264-x.
[58] Oliveira CS, Sigbjörnsson R, Ólafsson S. A comparative study on strong ground motion in two volcanic environments: Azores and Iceland. In: Proceedings of the proceedings of the 13th world conference on earthquake engineering. Citeseer; 2004. p. 13.
[59] Danciu L, Nandan S, Reyes C, Basili R, Weatherill G, Beauval C, Rovida A, Vilanova S, Sesetyan K, Bard PY, Cotton F, Wiemer S, Giardini D. The 2020 update of the European seismic hazard model: model overview. EFEHR Techn Rep 001 2021. https://doi.org/10.12686/a15. v1.0.0.
[60] Danciu L, Giardini D, Weatherill G, Basili R, Nandan S, Rovida A, Beauval C, Bard PY, Pagani M, Reyes CG, Sesetyan K, Vilanova S, Cotton, Wiemer S. The 2020 European seismic hazard model: overview and results. Nat Hazards Earth Syst Sci 2024;24(9):3049–73.
[61] Molkenthin C, Scherbaum F, Griewank A, Kuehn N, Stafford PJ. A Study of the sensitivity of response spectral amplitudes on seismological parameters using algorithmic differentiation. Bull Seismol Soc Am 2014;104(5):2240–52.
[62] Douglas J, Jousset P. Modeling the difference in ground-motion magnitude scaling in small and large earthquakes. Seismol Res Lett 2011;82(4):504–8.
[63] Atkinson GM, Silva W. An empirical study of earthquake source spectra for California earthquakes. Bull Seismol Soc Am 1997;87(1):97–113.
[64] Bora SS, Cotton F, Scherbaum F, Edwards B, Traversa P. Stochastic source, path and site attenuation parameters and associated variabilities for shallow crustal European earthquakes. Bull Earthq Eng 2017;15:4531–61.
[65] Chiou BJ, Youngs RR. An NGA model for the average horizontal component of peak ground motion and response spectra. Earthq Spectra 2008;24(1):173–215.
[66] Kotha SR, Cotton F, Bindi D. Empirical models of shear-wave radiation pattern derived from large datasets of ground-shaking observations. Sci Rep 2019;9:1–11.
[67] Ambeh WB, Fairhead JD. Coda Q estimates in the mount Cameroon volcanic region, West Africa. Bull Seismol Soc Am 1989;79(5):1589–600.
[68] Mayeda K, Koyanagi S, Aki K. Site amplification from S-wave coda in the Long Valley Caldera region, California. Bull Seismol Soc Am 1991;81(6):2194–213.
[69] Londono JM. Temporal change in coda Q at Nevado del Ruiz Volcano, Colombia. J Volcanol Geoth Res 1996;73(129):139.
[70] Olsen KB, Mayhew JE. Goodness-of-fit criteria for broadband synthetic seismograms, with application to the 2008 Mw 5.4 Chino Hills, California, earthquake. Seismol Res Lett 2010;81(5):715–23.
[71] Escuer M, Oliveira CS, Dessai P, Lopes H. Study of the amplification of seismic waves in Monte das Moças, Horta. Confrontation of records obtained at the seismographic and accelerographic stations. In: Proceedings 2 Simpo'sio de Meteorologia e Geofísica da APMG, 3 Encontro Luso-Espanhol de Meteorologia. Portugal: Évora; 2001. p. 105–9.
[72] García-Soto AD, Jaimes MA. Ground-motion prediction model for vertical response spectra from Mexican interplate earthquakes. Bull Seismol Soc Am 2017;107(2):887–900.
[73] Jaimes MA, García-Soto AD. Updated ground motion prediction model for Mexican intermediate-depth intraslab earthquakes including V/H ratios. Earthq Spectra 2020;36(3):1298–330.
[74] Jaimes MA, García-Soto AD. Ground-Motion duration prediction model from recorded Mexican interplate and intermediate-depth intraslab earthquakes. Bull Seismol Soc Am 2021;111(1):258–73.
[75] Jaimes MA, García-Soto AD. Vertical ground motion prediction equations for interplate and intermediate-depth intraslab earthquakes at Mexico city's hill zone. J Earthq Eng 2024;28(11):3038–67.
[76] Jaimes MA, García-Soto AD, Candia G. Horizontal and vertical ground-motion duration prediction models from interplate and intermediate-depth intraslab earthquakes in Mexico city. Bull Seismol Soc Am 2024;114(3):1695–716.
[77] Phung VB, Loh CH, Chao SH, Abrahamson NA. Ground motion prediction equation for Taiwan subduction zone earthquakes. Earthq Spectra 2020;36(3):1331–58.
[78] Zhao JX, Zhou S, Zhou J, Zhao C, Zhang H, Zhang Y, Irikura K. Ground-motion prediction equations for shallow crustal and upper-mantle earthquakes in Japan using site class and simple geometric attenuation functions. Bull Seismol Soc Am 2016;106(4):1552–69.
[79] Eurocode 8. Design of structures for earthquake resistance Part 1: general rules, seismic actions and rules for buildings. Brussels: European Committee for Standardization (CEN); 2004. EN 1998-1:2004.
[80] Seyhan E, Stewart JP. Semi-empirical nonlinear site amplification from NGA-West2 data and simulations. Earthq Spectra 2014;30(3):1241–56.