



# A far-field ground motion prediction model for interface earthquakes at the hill zone of Mexico City

M. Leonardo-Suárez<sup>1</sup> · A. F. Hernández<sup>2</sup> ·  
P. Quinde<sup>3</sup>

Received: 12 August 2022 / Accepted: 17 January 2023 / Published online: 6 February 2023  
© The Author(s), under exclusive licence to Springer Nature B.V. 2023

**Abstract** This study presents an updated attenuation model to predict the peak ground acceleration (PGA), peak ground velocity (PGV), 5% damped pseudo-spectral acceleration (SA), and the average spectral acceleration (AvgSA) at the hill zone of Mexico City for interface earthquakes. The strong-motion dataset comprises 33 earthquakes recorded at CU station, covering a moment magnitude ( $M_w$ ) range from 6.0 to 8.1 and a source-to-site distance ( $R_{rup}$ ) range from 240 to 490 km. Given the small number of available observations, a Bayesian regression scheme is used to obtain the coefficients of the ground-motion prediction model (GMPM). In addition, the epistemic uncertainty in the estimation of the regression coefficients is evaluated, showing its impact on the framework of a probabilistic seismic hazard analysis

(PSHA). The results are compared with models previously developed for the CU station, discussing the differences observed between the median predictions and their standard deviations. Likewise, seismic hazard curves are computed and compared with empirical curves obtained by counting the number of times per year that a given value of ground-motion intensity is exceeded. The results show that the dispersion of the GMPM proposed is lower than the previous models for PGA and SA, which means better predictability and more reliable estimates of the seismic hazard at the site.

**Keywords** Ground-motion prediction model · Single-station dataset · Bayesian regression · Average spectral acceleration · Intensity measures

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10950-023-10132-0>.

M. Leonardo-Suárez  
Instituto de Ingeniería, Universidad Nacional Autónoma de México, Mexico City, Mexico  
e-mail: MLeonardoS@iingen.unam.mx

A. F. Hernández  
ISTerre, Université Grenoble Alpes, Grenoble, France  
e-mail: andres.hernandez@univ-grenoble-alpes.fr

P. Quinde (✉)  
Facultad de Ciencia Y Tecnología, Universidad del Azuay,  
Cuenca, Ecuador  
e-mail: pabloquinde@uazuay.edu.ec

## 1 Introduction

Mexico City, located more than 280 km inland from Mexico's Pacific Coast, has been affected by several interface earthquakes throughout its history, which have caused a considerable number of casualties and severe damage to the civil infrastructure. Some examples of these major events are (1) the 1957 San Marcos earthquake ( $M$  7.5) that damaged more than 1000 buildings and claimed the life of 39 people (Orozco and Reinoso 2007); (2) the 1979 Petatlán earthquake ( $M_w$  7.4) that damaged more than 50 structures and caused the collapse of three buildings (Alonso et al.

1979); and (3) the 1985 Michoacán earthquake ( $M_w$  8.1) that produced the collapse of 256 buildings and more than 10,000 deaths (Astiz et al. 1987), being the latter the most destructive earthquake in Mexico City's history whose economic losses amounted to US\$ 4 billion (at 1985 value) according to Mexican Insurance Association. The damage pattern observed in these interface events was attributed to the well-known local site effects in the Mexico City basin (MCB) (Meli and Miranda 1985). The ground motion amplification and duration at MCB are mainly attributed to the elastic behavior of the clayed soil deposits (Díaz-Rodríguez 1989; Díaz-Rodríguez 1992; Romo and Ovando-Shelley 1996; Taboada-Urtuzuástegui et al. 1999; Díaz-Rodríguez and Santamarina 2001) whose parameters, shear modulus and damping ratio, do not show a significant reduction at high values of shear strain. This condition amplifies seismic waves in these deposits (Mayoral et al. 2008, 2016, 2019).

In the early 1960s, the Mexico City Accelerometric Network (MCAN) began operating to study the dynamic amplification of the seismic waves in the lakebed deposits. Nowadays, the MCAN has over 149 strong-motion stations distributed throughout the city at three different geotechnical zones: zone I (*hill*), formed by volcanic tuffs and lava flows; zone II (*transition*), composed of 20 m sand and silt layers interbedded with lacustrine clay layers; and zone III (*lakebed*), consisting of 30–80-m clay deposit highly compressible interbedded with sand layers (Mayoral et al. 2019). The station Ciudad Universitaria (herein, CU), located within the *hill zone* of the National Autonomous University of Mexico (UNAM, by its Spanish acronym), is one of the most important and oldest stations of the MCAN. It has been in continuous operation since 1962 and has become the reference station to study the site effect at the *lakebed zone* (Ordaz et al. 1988; Singh et al. 1988a; Lermo and Chávez-García 1993; Reinoso and Ordaz 1999; Montalvo-Arrieta et al. 2002). Furthermore, its information has helped develop and calibrate the early earthquake damage assessment system available for the management of seismic risk in the city (Iglesias et al. 2007; Ordaz et al. 2017).

Since the 1980s, several studies have been conducted to predict the ground motion at the *hill zone* of Mexico City from interface earthquakes recorded at the station CU. For instance, Singh et al. (1987) used a set of 16 earthquakes ( $5.6 \leq M_s \leq 8.1$  and

$282 \leq R_{rup} \leq 466$ ) recorded at this station to derive a ground-motion prediction model (GMPM) in terms of the peak ground acceleration (PGA) and velocity (PGV). Based on the same strong-motion dataset, Castro et al. (1988) developed a further GMPM to estimate the Fourier amplitude spectra (FAS) of the ground acceleration for frequencies between 0.4 and 5.0 Hz. This last model was improved by Ordaz et al. (1994) using more records and applying a Bayesian regression approach instead of the multiple linear regression (MLR) method used in the previous works. Years later, Reyes (1999) employed 17 earthquakes ( $6.1 \leq M_w \leq 8.1$  and  $280 \leq R_{rup} \leq 466$ ) to derive a new GMPM for PGA and the 5% damped pseudo-spectral acceleration (SA) for 60 structural periods between 0.1 and 6.0 s applying the Bayesian regression approach proposed by Ordaz et al. (1994). This latest is updated by Jaimes et al. (2006) adding to the dataset 5 events that occurred between 1996 and 2004 ( $6.0 \leq M_w \leq 7.5$  and  $301 \leq R_{rup} \leq 526$ ) and deriving the GMPM in terms of the PGA and SA for 30 structural periods between 0.2 and 6.0 s. These latest two GMPMs have been used over time to perform site-specific probabilistic seismic hazard analysis (PSHA).

In particular, the GMPMs proposed by Reyes (1999) and Jaimes et al. (2006) (herein referred as to R99 and J06, respectively) have been used in the current Mexico City Standard for Seismic Design (NTC-DS 2017) to define the design spectrum at the *hill zone* based on the uniform hazard spectrum (UHS) obtained at CU for 250-year return period. This UHS has been used as a reference to estimate seismic hazard and design spectra at other city sites through the spectral amplification functions pre-computed between soft sites and the reference station CU (Ordaz 2017). Despite the good performance of R99 and J06, in terms of their predictive capability of ground motion for Mexican interface earthquakes, it is essential to update the database to consider the moderate earthquakes ( $6.0 < M_w < 7.4$ ) that have occurred during 2004–2021. Furthermore, it is necessary to evaluate other intensity measures (IMs) that better represent the ground motion's severity and be a good predictor of the structural response.

This study presents an updated GMPM for the *hill zone* in Mexico City, based on a dataset from interface earthquakes recorded at station CU with a moment magnitude range ( $M_w$ ) from 6.0 to 8.1 and a source-to-site distance range ( $R_{rup}$ ) from 240

to 490 km. Four IMs are used to characterize the ground motion: peak ground acceleration (PGA), peak ground velocity (PGV), 5% damped pseudo-spectral acceleration (SA), and the average spectral acceleration (AvgSA). The results are compared with previously derived GMPMs for SA, discussing the differences observed between the predicted median response spectra and their residual distribution. Finally, the GMPM derived is used to compute seismic hazard curves and compare the results with empirical curves obtained directly from counting the times a given IM value has been exceeded per year. The seismic hazard curve *directly* estimated from GMPM-AvgSA is compared with the exceedance rate of AvgSA values derived *indirectly* from existing GMPM-SAs (R99 and J06) using a consistent SA inter-period correlation model proposed by Rodríguez-Castellanos et al. (2021).

## 2 Strong-motion dataset

The strong-motion dataset used in this study corresponds to the Mexican interface earthquakes recorded at the station CU, located within the *hill zone* of Mexico City. The dataset includes 33 record pairs (two horizontal components) from 33 reverse-faulting events with  $M_w > 6$ , which occurred between 1965 and 2021 along Mexico's Pacific Coast. The small events ( $M_w < 6$ ) were excluded from the dataset because their ground-motion intensity levels (e.g.,  $PGA < 1 \text{ cm/s}^2$ ) are not representative of most engineering interests. Table 1 summarizes the seismic source parameters (i.e., location, magnitude, and focal mechanism) of the events analyzed. These parameters were taken from the recently published earthquake catalog by Sawires et al. (2019), which includes an extensive and careful review of all Poissonian independent earthquakes that occurred in Mexico during a period ranging from 1787 to 2018. The source parameters of the events that occurred after 2018 were obtained from the earthquake reports published by the Servicio Sismológico Nacional (SSN (Mexican National Seismological Service)) through its website (see SSN (2022)). Figure 1 shows the epicenters and focal mechanism of the dataset used, covering a magnitude range from 6.0 to 8.1 and a source-to-site distance range from 240 to 490 km.

### 2.1 Local site conditions

According to the geotechnical zoning of the NTCDS (2017), the *hill zone* is formed by rocks and hard soils where sand deposits may be interbedded with volcanic flows. Based on this definition, it could be assumed that the local site conditions of CU correspond to *firm soil*. However, several studies carried out at this site (Singh et al. 1988b, 1995; Ordaz and Singh 1992; Montalvo-Arrieta et al. 2002, 2003) have shown the existence of a large ground motion amplification in a range of frequencies between 0.2 and 10 Hz. Montalvo-Arrieta et al. (2003) report that the spectral ratios of stations located in the *southwestern hill zone* (including CU) exhibit relative amplifications up to 4 times larger than stations located in the *northern* for frequencies between 0.2 and 3 Hz. Ordaz and Singh 1992 suggest that this amplification can be related to the shallow (< 1 km) sediments that lie below the volcanic rocks that cover the *hill zone* of the Valley of Mexico. As regards, several studies have been conducted near station CU to determine the shear wave velocity profile (see Singh et al. 1995; Kagawa 1996; Flores Estrella and Aguirre González 2003) through geophysical prospecting techniques, whose results indicate that the site can be classified as a *soft rock* with an average shear wave velocity in the upper 30 m ( $V_s 30$ ) of 450 to 760 m/s.

### 2.2 Source-to-site distance

The source-to-site distance selected to develop the GMPM was the closest to the rupture plane ( $R_{rup}$ ). This distance was computed following the procedure presented by Chiou and Youngs (2008), simulating the rupture dimensions for each event through the scaling relationship derived by Ramírez-Gaytán et al. (2014) for Mexican subduction earthquakes. For eight of the most significant events ( $M_w > 7$ ) in the dataset, the distance was inferred directly from the finite-fault geometry obtained by kinematic inversion, whose information was taken from the *SRCMOD* database compiled by Mai and Thingbaijam (2014).

Figure 2 shows the magnitude-distance (left panel) and magnitude-focal depth (right panel) distributions for the compiled dataset (hereafter, CU21). As can be seen, CU21 covers a magnitude range from 6.0 to 8.1 and a source-to-site distance range from 240 to 490 km. Something to highlight in this figure is the

**Table 1** Reverse-faulting interface events with  $M_w > 6$  recorded at CU station

ID <sup>a</sup>	Date (yyyy/mm/dd)	Lat (°N)	Lon (°W)	Depth (km)	Mag ( $M_w$ )	$M_o$ (dyne-cm)	Focal mechanism			$R_{rup}$ (km)
							$\theta$	$\delta$	$\lambda$ <sup>b</sup>	
1	1965/08/23	16.300	95.800	16.0	7.50	1.9953E+27	268	14	54	432.13
2	1968/08/02	16.590	97.700	16.0	7.30	1.0000E+27	287	12	76	293.62
3	1976/06/07	17.400	100.640	29.0	6.40	5.4000E+25	289	18	83	249.15
4	1978/03/19	17.030	99.740	360	6.60	8.8700E+25	279	10	102	247.83
5	1978/11/29	16.000	96.690	18.0	7.70	5.2700E+27	274	7	57	382.75
6*	1979/03/14	17.460	101.460	20.0	7.40	1.3700E+27	293	14	90	270.24
7*	1981/10/25	17.750	102.250	20.0	7.25	8.4900E+26	300	14	90	323.96
8	1982/06/07	16.350	98.370	20.0	6.90	2.9000E+26	268	10	48	315.88
9	1982/06/07	16.400	98.540	15.0	7.00	2.6600E+26	286	12	76	298.72
10*	1985/09/19	18.140	102.710	16.0	8.10	1.1500E+28	300	14	61	301.15
11*	1985/09/21	17.600	101.800	20.0	7.42	1.5300E+27	300	14	100	287.80
12	1989/04/25	16.795	99.275	23.0	6.90	2.3900E+26	276	10	66	259.94
13	1993/05/15	16.430	98.740	20.0	6.00	1.4100E+25	242	35	30	314.99
14	1993/10/24	16.540	98.980	19.0	6.60	1.0100E+26	276	17	67	291.99
15*	1995/09/14	16.752	98.667	21.0	7.30	8.9000E+26	289	11	75	248.55
16	1996/07/15	17.293	101.241	18.7	6.60	9.9500E+25	297	21	93	295.33
17	1998/02/03	15.690	96.370	24.0	6.30	3.7200E+25	288	42	104	486.34
18	2000/08/09	17.990	102.660	16.0	6.50	7.1300E+25	291	43	91	376.66
19*	2003/01/22	18.540	104.270	16.0	7.50	2.0200E+27	300	22	93	484.81
20	2004/01/01	17.340	101.420	15.0	6.00	1.4600E+25	299	13	92	311.88
21	2010/06/30	16.290	97.959	27.8	6.30	3.0600E+25	286	12	72	344.05
22*	2012/03/20	16.246	98.457	18.0	7.42	1.7800E+27	296	13	91	294.99
23	2012/04/11	17.921	103.068	20.0	6.70	1.2000E+26	282	25	77	413.08
24	2013/04/22	17.856	102.280	26.6	6.10	1.6600E+25	113	29	77	357.76
25	2013/08/21	16.857	99.519	18.9	6.20	2.1840E+25	287	19	74	264.36
26	2014/04/18	17.370	101.050	18.0	7.25	9.9762E+26	302	20	99	263.67
27	2014/05/08	16.986	100.916	10.0	6.50	4.4110E+25	315	26	99	302.23
28	2014/05/10	17.036	100.893	10.0	6.10	1.2490E+25	313	21	105	301.05
29	2016/05/08	16.323	97.877	22.0	6.00	7.0320E+24	284	19	66	348.20
30*	2018/02/16	16.218	98.013	19.5	7.20	7.1465E+26	297	12	91	322.84
31	2018/02/19	16.247	97.775	10.2	6.00	8.3290E+24	285	23	84	359.61
32 <sup>§</sup>	2020/06/23	15.784	96.120	22.0	7.41	1.6400E+27	266	17	60	456.84
33 <sup>§</sup>	2021/09/07	16.820	99.780	10.0	7.10	5.0119E+26	288	13	89	261.02

<sup>a</sup>Earthquake ID. The symbols represent the data source from which the finite-fault geometry information was obtained (i.e., \*Mai and Thingbaijam (2014); <sup>§</sup> National Seismological Service (SSN 2022))

<sup>b</sup>The variables in the parentheses represent  $\theta$  strike,  $\delta$  dip, and  $\lambda$  rake

low sample size of the dataset for magnitudes and distances beyond 7.5 and 350 km, respectively. This lack of information can be compensated through prior theoretical knowledge about certain seismological parameters related to the scaling of the ground motion intensity with the magnitude and distance, whose *prior information* can be formally integrated into the regression process using a Bayesian approach, obtaining reasonable results as it is shown by Ordaz et al. 1994; Jaimes et al. 2006; Wang and Takada 2009.

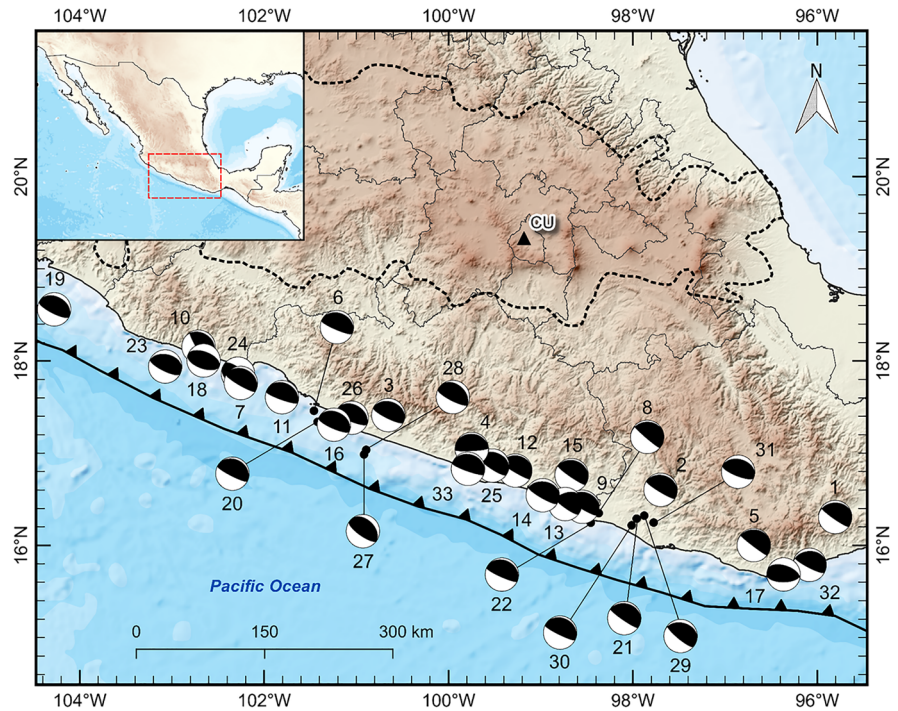
Notably, the last-mentioned work indicates that using *prior information* in a Bayesian approach is particularly effective when the dataset is small.

### 3 Ground-motion intensity measures

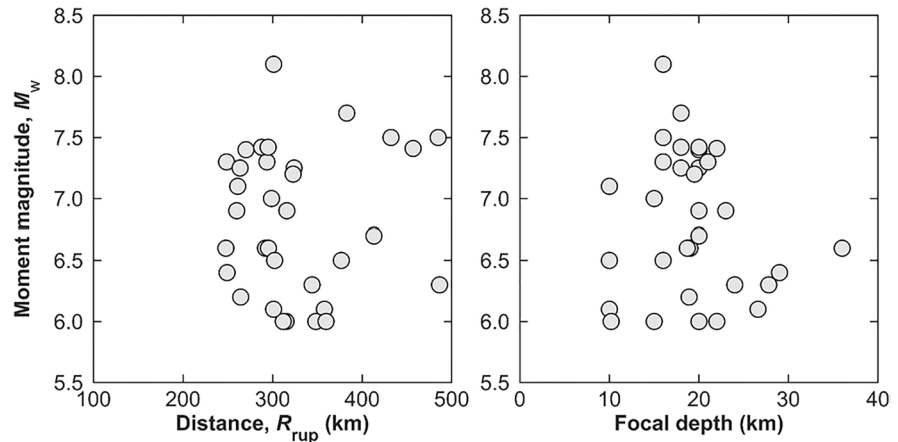
The intensity measures (IMs) commonly used to derive a GMPM are peak ground acceleration (PGA), peak ground velocity (PGV), and spectral



**Fig. 1** Location of the epicenters and focal mechanisms of earthquakes recorded at station CU (filled triangle) in Mexico City. The dashed line represents the boundary of the Trans-Mexican volcanic belt (TMBV). The epicentral data and source parameters are listed in Table 1



**Fig. 2** Magnitude-distance and magnitude-depth distributions of the dataset used in the regression analysis



acceleration (SA). However, several studies (Giovenale et al. 2004; Luco and Cornell 2007; Eads et al. 2015; Kohrangi et al. 2016a, b) have demonstrated that these IMs used in performance-based assessments can be *inefficient* and *insufficient* to predict the structural demand, especially for tall or long-period buildings whose higher modes contribute significantly to the structural response. This condition has led to the development of new IMs better correlated with the structural response of nonlinear systems, such as the average spectral acceleration (AvgSA).

Particularly, Eads et al. (2015) define AvgSA as the geometric mean of SA ordinates in the range of  $0.2T_1-3T_1$ , where  $T_1$  is the fundamental period of vibration of the structure. In Eads et al. (2015), the authors used almost 700 buildings of different heights (4 to 16 stories) and structural systems (moment-resist frame and shear wall structures) to evaluate the *efficiency* and *sufficiency* of AvgSA as IM to predict the collapse intensities. Their results suggest that AvgSA is a better option than SA since it reduces the variability of the predicted structural demand and the

IM dependency with the ground-motion characteristics, such as magnitude ( $M$ ), source-to-site distance ( $R$ ), and fault mechanism.

On the other hand, Kohrangi et al. (2017) mention that selecting the upper period limit ( $T^*$ ) to compute AvgSA is not a trivial matter since it depends on the objective of the analysis. In that study, the authors recommend a value of  $T^*$  equal to  $1.5 T_1$  or  $2 T_1$  to predict the peak inter-story drift ratio (IDR) based on the results obtained in (Kohrangi et al. 2016a, b), in which the nonlinear response of three reinforced concrete buildings (3, 5, and 8 stories) subjected to both horizontal components of the ground motion are evaluated. Its results show that using  $T^* = 1.5 T_1$  gets the lowest dispersion in response estimation of the IDR, which engineering demand parameter (EDP) correlates well with the collapse state. Likewise, they show the *efficiency* and *sufficiency* of predicting the IDR using AvgSA [ $0.2 T_1 - 1.5 T_1$ ].

Based on the above, it was decided to include the AvgSA as an IM to derive the GMPM. The intention is not to replace SA as IM to define design loads in the standard for seismic design but rather to have an IM better correlated with the nonlinear structural behavior that allows for obtaining more accurate seismic risk estimations. Therefore, it was decided to keep the typical IMs used in previous works to derive the GMPM such as PGV, PGA, and SA, which metrics have been used with different objectives over time (e.g., seismic risk assessment, calibration of the early warning system, obtaining the design spectrum, among others).

In this study, the AvgSA is estimated for 40 periods,  $T_1$ , equally spaced between 0.1 s and 4 s, considering a period range of  $0.1 T_1 - 1.5 T_1$  and a uniform spacing ( $\Delta$ ) of 0.1 in the averaging process through the Eq. (1). This period range corresponds to the one proposed in the NTCDS (2017) to evaluate the nonlinear response of tall buildings in Mexico City. Therefore, the *GMPM-AvgSA* derived here could be used to obtain the conditional mean spectrum (CMS) and select acceleration time histories that can be used in dynamic analysis.

$$\text{AvgSA}(T_1) = \left[ \prod_{i=1}^N \text{SA}(C_i T_1) \right]^{1/N} \quad (1)$$

where SA stands the 5% damped spectral acceleration value,  $N$  is the number of periods used to compute AvgSA, in this case,  $N=15$ ;  $C_i$  is the  $i$ th coefficient of the period range used in the averaging process,  $C_i = i\Delta$  and  $T_1$  is the fundamental period of vibration of the structure.

Likewise, three IMs were used to characterize the ground motion: PGV, PGA, and SA. For these IMs, the quadratic mean (QM) of the maximum response of both horizontal components was computed to derive the GMPM, that is,  $\text{SA}_{\text{QM}} = [(\text{SA}_N^2 + \text{SA}_E^2) / 2]^{0.5}$ . The QM was selected to be consistent with the definition adopted in other GMPMs developed for Mexico City (e.g., Reyes 1999; Jaimes et al. 2015) and used in the PSHA. For SA, the 5% damped response spectra at 102 structural periods between 0.05 s and 6 s were computed. The processing of the acceleration time histories used in the analysis is described as follows.

## 4 Data processing and analysis

### 4.1 Data processing

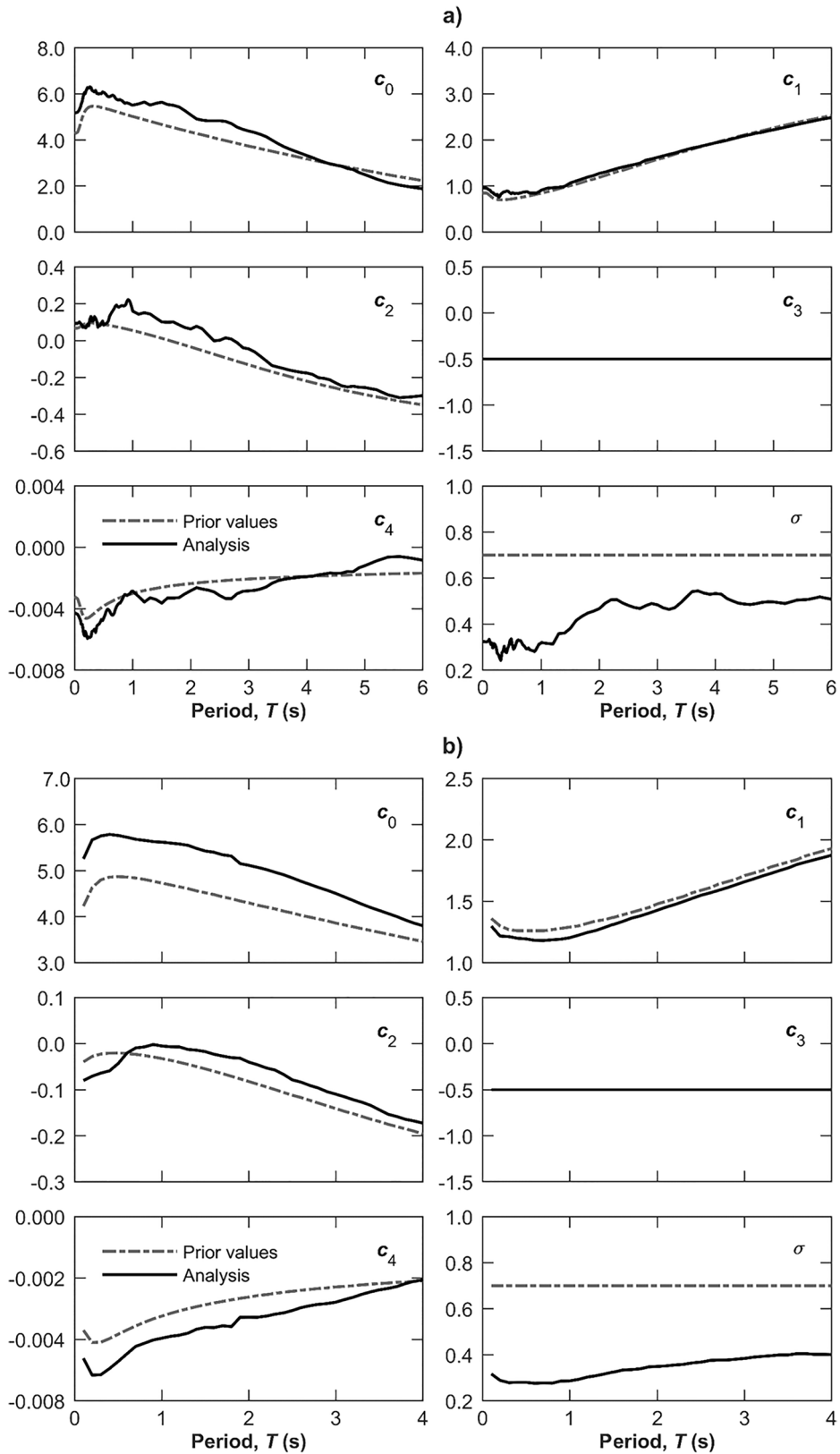
The records were processed applying a baseline correction and high-pass filter using two values of corner frequency ( $f_{\text{hp}}$ ) depending on the type and resolution of the recording instrument: (1)  $f_{\text{hp}} = 0.30$  for the events recorded at analog and digital accelerographs with low resolution ( $< 16$  bit) and flat response in the range of 0.1 to 35 Hz and (2)  $f_{\text{hp}} = 0.05$  for the events recorded at modern digital accelerographs. The minimum usable oscillator frequency ( $f_o$ ) to compute SA was determined for each record as 1.25 times the corner frequency of the high-pass filter (Abrahamson and Silva 1997; Ancheta et al. 2014; Bindi et al. 2014; Kohrangi et al. 2018). This condition implies that for AvgSA the lowest usable  $f_o$  on the data selection is constrained by the upper limit of the period range  $0.1 T_1 - 1.5 T_1$  selected, which corresponds to  $1.875 f_{\text{hp}}$ . Thus, the number of record pairs in the final dataset varies as a function of the oscillator period,  $T$  (where  $T = 1/f_o$ ), including a maximum of 33 strong-motion records to estimate SA and AvgSA for a  $T \leq 2.7$  s and  $T \leq 1.8$  s, respectively, and 24 to evaluate the response for both IMs at longer periods.

The PGA value was read directly from each record processed, and the 5% damped SA for 102 structural

**Table 2** Coefficients obtained (see Eq. 2)

<i>T</i> (s)	SA ( <i>T</i> ) <sup>a</sup>					AvgSA ( <i>T</i> ) <sup>a</sup>				
	<i>c</i> <sub>0</sub>	<i>c</i> <sub>1</sub>	<i>c</i> <sub>2</sub>	<i>c</i> <sub>4</sub>	$\sigma$	<i>c</i> <sub>0</sub>	<i>c</i> <sub>1</sub>	<i>c</i> <sub>2</sub>	<i>c</i> <sub>4</sub>	$\sigma$
PGV	2.7281	1.4706	-0.0434	-0.0021	0.4286					
PGA	5.1671	0.9599	0.0925	-0.0043	0.3248					
0.05	5.2099	0.9543	0.0944	-0.0044	0.3225					
0.10	5.4459	0.9146	0.0946	-0.0048	0.3335	5.2567	1.2978	-0.0800	-0.0046	0.3170
0.20	6.1631	0.8212	0.0727	-0.0057	0.3115	5.6687	1.2175	-0.0706	-0.0052	0.2892
0.30	6.1508	0.8243	0.1053	-0.0055	0.2426	5.7540	1.2128	-0.0639	-0.0051	0.2793
0.40	6.0079	0.8925	0.0730	-0.0050	0.2759	5.7877	1.1997	-0.0590	-0.0049	0.2806
0.50	6.0645	0.8560	0.0831	-0.0046	0.3089	5.7642	1.1942	-0.0429	-0.0047	0.2806
0.60	5.8783	0.8368	0.1315	-0.0041	0.3310	5.7265	1.1821	-0.0204	-0.0045	0.2763
0.70	5.8213	0.8693	0.1874	-0.0040	0.3063	5.6818	1.1806	-0.0093	-0.0042	0.2781
0.80	5.7247	0.8557	0.1884	-0.0035	0.2885	5.6589	1.1871	-0.0070	-0.0041	0.2771
0.90	5.6159	0.8464	0.2132	-0.0031	0.2856	5.6316	1.1921	-0.0020	-0.0040	0.2864
1.00	5.5138	0.9118	0.1594	-0.0028	0.3198	5.6186	1.2039	-0.0050	-0.0040	0.2869
1.10	5.5682	0.9517	0.1434	-0.0032	0.3162	5.6002	1.2242	-0.0068	-0.0039	0.2933
1.20	5.6461	0.9630	0.1513	-0.0036	0.3139	5.5799	1.2460	-0.0070	-0.0039	0.3033
1.30	5.5169	0.9728	0.1441	-0.0033	0.3589	5.5510	1.2627	-0.0126	-0.0038	0.3092
1.40	5.5904	0.9919	0.1222	-0.0034	0.3613	5.4831	1.2873	-0.0142	-0.0037	0.3171
1.50	5.6398	1.0678	0.1012	-0.0036	0.3819	5.4274	1.3130	-0.0173	-0.0036	0.3255
1.60	5.5405	1.1104	0.0993	-0.0033	0.4162	5.3988	1.3365	-0.0231	-0.0036	0.3340
1.70	5.5055	1.1516	0.1010	-0.0033	0.4340	5.3445	1.3629	-0.0270	-0.0036	0.3360
1.80	5.4324	1.1951	0.0851	-0.0033	0.4481	5.3166	1.3812	-0.0304	-0.0036	0.3393
1.90	5.2698	1.2280	0.0707	-0.0030	0.4566	5.1532	1.4037	-0.0317	-0.0033	0.3477
2.00	5.0881	1.2747	0.0629	-0.0028	0.4676	5.1154	1.4284	-0.0401	-0.0033	0.3498
2.20	4.8606	1.3411	0.0635	-0.0027	0.5058	5.0254	1.4749	-0.0527	-0.0032	0.3561
2.40	4.8331	1.4113	-0.0015	-0.0029	0.4919	4.9666	1.4991	-0.0572	-0.0032	0.3588
2.60	4.8047	1.4658	0.0149	-0.0033	0.4726	4.9109	1.5246	-0.0669	-0.0031	0.3651
2.80	4.5538	1.5624	-0.0117	-0.0030	0.4802	4.8426	1.5496	-0.0786	-0.0031	0.3684
3.00	4.3871	1.6209	-0.0452	-0.0028	0.4840	4.7690	1.5684	-0.0835	-0.0030	0.3755
3.20	4.2437	1.6869	-0.0850	-0.0027	0.4641	4.3509	1.7058	-0.1216	-0.0026	0.3934
3.40	3.9932	1.7560	-0.1353	-0.0023	0.4927	4.2715	1.7273	-0.1271	-0.0025	0.3959
3.60	3.7148	1.8239	-0.1515	-0.0021	0.5414	4.1992	1.7507	-0.1340	-0.0024	0.3999
3.80	3.5121	1.8791	-0.1664	-0.0020	0.5350	4.1349	1.7721	-0.1437	-0.0024	0.3998
4.00	3.3245	1.9383	-0.1759	-0.0019	0.5326	4.0692	1.8000	-0.1532	-0.0023	0.4049
4.20	3.1290	1.9937	-0.2012	-0.0018	0.5110					
4.40	2.9587	2.0580	-0.2130	-0.0017	0.4908					
4.60	2.8559	2.1169	-0.2451	-0.0017	0.4852					
4.80	2.7152	2.1657	-0.2475	-0.0016	0.4944					
5.00	2.4967	2.2212	-0.2546	-0.0012	0.4964					
5.20	2.3061	2.2762	-0.2666	-0.0009	0.4896					
5.40	2.1444	2.3345	-0.2942	-0.0006	0.5040					
5.60	2.0383	2.3891	-0.3098	-0.0006	0.5074					
5.80	1.9543	2.4432	-0.3053	-0.0007	0.5182					
6.00	1.8801	2.4877	-0.2985	-0.0008	0.5080					

<sup>a</sup>The period-independent coefficient used in the regression analysis for all IMs:  $c_3 = -0.5$ . Note that the median values obtained from these coefficients are expressed in acceleration units (cm/s<sup>2</sup>) for PGA, SA, and AvgSA. For PGV, the results are expressed in velocity units (cm/s).



◀**Fig. 3** Regression coefficients of the GMPM for **a** SA and **b** AvgSA. The solid and dashed lines represent the posterior and prior coefficients, respectively, used in this study

periods between 0.05 s and 6 s was computed. The PGV values were obtained by integrating the acceleration-time histories after applying the baseline correction and high-pass filter. For these IMs, the quadratic mean of the maximum response of both horizontal components was computed to derive the GMPM. Finally, the AvgSA values were estimated for 40 periods,  $T_1$ , equally spaced between 0.1 s and 4 s using Eq. (1).

#### 4.2 Regression analysis

The regression of the dataset was performed using the Bayesian approach proposed by Ordaz et al. (1994), which was considered adequate to obtain the GMPM given the small number of available observations. Besides, it is not possible to apply a two-step or random-effect approach because the observation comes from a *single station*. The Bayesian approach makes it possible to incorporate theoretical knowledge about the phenomenon being studied to constrain the values of some seismological parameters that cannot be directly derived from the dataset. This *prior information*, combined with the empirical data through the Bayes theorem, provides a rational and stable numerical solution, as shown by Ordaz et al. (1994), Jaimes et al. (2006), and Wang and Takada (2009). The functional form adopted in the regression analysis to develop the GMPM is shown in Eq. (2).

$$\ln Y = c_0 + c_1(M_w - 6) + c_2(M_w - 6)^2 + c_3 \ln R + c_4 R + \varepsilon \tag{2}$$

where  $Y$  is the predicted median of a ground-motion parameter: PGA, PGV, SA, or AvgSA. The predicted values for acceleration and velocity are in  $\text{cm/s}^2$  and  $\text{cm/s}$ , respectively.  $M_w$  is the moment magnitude;  $R_{\text{rup}}$  is the closest distance to the rupture plane in kilometers;  $c_i$  are the coefficients to be determined by the regression analysis, and  $\varepsilon$  is the random error normally distributed with zero mean and variance  $\sigma_\varepsilon^2$ . In this case, being a *single-station* analysis,  $\sigma_\varepsilon^2$  represents the variance of the interevent residuals (Atkinson 2006). Note that SA and AvgSA coefficients depend on the fundamental vibration period of the structure,  $T$ ; however, for simplicity, this term has been removed from Eq. (2).

The *prior information* is an intrinsic part of a Bayesian approach and reflects our state of knowledge about the phenomenon analyzed prior to the observation and analysis of data. In this study, the *prior* probability distributions of the coefficients,  $c_i$ , are obtained from the ground-motion intensity (GMI) values estimated theoretically at CU without using the strong-motion recordings. For that purpose, a far-field point source model proposed by Singh et al. (1989) and the random vibration theory (RVT) was used to compute the Fourier amplitude spectrum (FAS) and predict the *GMI* in terms of PGA, PGV, and SA. The AvgSAs were obtained from the SAs using Eq. (1). The frequency-dependent attenuation parameters required to compute the FAS were taken from Ordaz and Singh (1992). Likewise, the strong-motion duration parameter used in the RVT method (Liu and Pezeshk 1999) was computed through the relationship proposed by López-Castañeda and Reinoso (2021). The expected values,  $E' [c_i]$ , and standard deviations,  $\sigma' [c_i]$ , of *prior* coefficients, were established following the procedure suggested by Reyes (1999), whose steps are briefly described below.

1. The *prior* expected values of  $c_0$ ,  $c_1$ , and  $c_2$  were estimated by applying a least squared fit to *GMIs* predicted for a specific source-to-site distance ( $R_{\text{rup}}=290$  km) and a magnitude ( $M_w$ ) range between 6.0 and 8.2. This procedure is considered adequate since, as shown in Eq. (2), these coefficients are independent of the distance;  $c_0$  represents the site-effects amplifications, while  $c_1$  and  $c_2$  control the magnitude scaling. The standard deviation of  $c_1$  and  $c_2$  was assigned according by Ordaz et al. (1994), adopting  $\sigma' [c_i]=E' [c_i] / 1.7$ .
2. The *prior* expected value of  $c_3$  was fixed at -0.5 for all periods. This value is consistent with the geometrical attenuation of surface waves at source-to-site distances greater than 100 km (Ordaz et al. 1992). A low standard deviation was assigned ( $\sigma' [c_3] \propto 0$ ) to ensure that the value remained constant after incorporating the observations.
3. The *prior* expected value of  $c_4$ , related to the quality factor  $Q$  (anelastic attenuation), was obtained as follows: (1) a magnitude value was set,  $M_w=7.0$ ; (2) for this value, the *GMI* was



**Table 3** Covariance matrixes obtained (see Eq. 5)

SA (T)						AvgSA (T)					
<b>PGV</b>	$c_0$	$c_1$	$c_2$	$c_3$	$c_4$						
$c_0$	556.488	-4.220	2.006	-114.572	0.325						
$c_1$	-4.220	0.277	-0.145	0.848	-0.002						
$c_2$	2.006	-0.145	0.083	-0.403	0.001						
$c_3$	-114.572	0.848	-0.403	23.599	-0.067						
$c_4$	0.325	-0.002	0.001	-0.067	0.00019						
<b>PGA</b>	$c_0$	$c_1$	$c_2$	$c_3$	$c_4$	<b><math>T=0.1</math></b>	$c_0$	$c_1$	$c_2$	$c_3$	$c_4$
$c_0$	319.506	-2.423	1.152	-65.781	0.186	$c_0$	305.722	-2.318	1.102	-62.943	0.178
$c_1$	-2.423	0.159	-0.083	0.487	-0.001	$c_1$	-2.318	0.152	-0.080	0.466	-0.001
$c_2$	1.152	-0.083	0.048	-0.231	0.001	$c_2$	1.102	-0.080	0.046	-0.221	0.001
$c_3$	-65.781	0.487	-0.231	13.549	-0.038	$c_3$	-62.943	0.466	-0.221	12.965	-0.037
$c_4$	0.186	-0.001	0.001	-0.038	0.00011	$c_4$	0.178	-0.001	0.001	-0.037	0.00011
<b><math>T=0.2</math></b>	$c_0$	$c_1$	$c_2$	$c_3$	$c_4$	<b><math>T=0.2</math></b>	$c_0$	$c_1$	$c_2$	$c_3$	$c_4$
$c_0$	293.782	-2.228	1.059	-60.485	0.171	$c_0$	253.350	-1.921	0.913	-52.161	0.148
$c_1$	-2.228	0.146	-0.077	0.448	-0.001	$c_1$	-1.921	0.126	-0.066	0.386	-0.001
$c_2$	1.059	-0.077	0.044	-0.213	0.001	$c_2$	0.913	-0.066	0.038	-0.183	0.00049
$c_3$	-60.485	0.448	-0.213	12.458	-0.035	$c_3$	-52.161	0.386	-0.183	10.744	-0.031
$c_4$	0.171	-0.001	0.001	-0.035	0.00010	$c_4$	0.148	-0.001	0.00049	-0.031	0.00009
<b><math>T=0.5</math></b>	$c_0$	$c_1$	$c_2$	$c_3$	$c_4$	<b><math>T=0.5</math></b>	$c_0$	$c_1$	$c_2$	$c_3$	$c_4$
$c_0$	288.499	-2.188	1.040	-59.397	0.168	$c_0$	236.724	-1.795	0.853	-48.738	0.138
$c_1$	-2.188	0.144	-0.075	0.440	-0.001	$c_1$	-1.795	0.118	-0.062	0.361	-0.001
$c_2$	1.040	-0.075	0.043	-0.209	0.001	$c_2$	0.853	-0.062	0.035	-0.171	0.00046
$c_3$	-59.397	0.440	-0.209	12.234	-0.035	$c_3$	-48.738	0.361	-0.171	10.039	-0.029
$c_4$	0.168	-0.001	0.001	-0.035	0.00010	$c_4$	0.138	-0.001	0.00046	-0.029	0.00008
<b><math>T=1.0</math></b>	$c_0$	$c_1$	$c_2$	$c_3$	$c_4$	<b><math>T=1.0</math></b>	$c_0$	$c_1$	$c_2$	$c_3$	$c_4$
$c_0$	309.823	-2.350	1.117	-63.788	0.181	$c_0$	247.918	-1.880	0.894	-51.043	0.145
$c_1$	-2.350	0.154	-0.081	0.472	-0.001	$c_1$	-1.880	0.123	-0.065	0.378	-0.001
$c_2$	1.117	-0.081	0.046	-0.224	0.001	$c_2$	0.894	-0.065	0.037	-0.180	0.00048
$c_3$	-63.788	0.472	-0.224	13.139	-0.037	$c_3$	-51.043	0.378	-0.180	10.513	-0.030
$c_4$	0.181	-0.001	0.001	-0.037	0.00011	$c_4$	0.145	-0.001	0.00048	-0.030	0.00009
<b><math>T=1.5</math></b>	$c_0$	$c_1$	$c_2$	$c_3$	$c_4$	<b><math>T=1.5</math></b>	$c_0$	$c_1$	$c_2$	$c_3$	$c_4$
$c_0$	443.511	-3.363	1.599	-91.312	0.259	$c_0$	319.290	-2.421	1.151	-65.737	0.186
$c_1$	-3.363	0.221	-0.116	0.676	-0.002	$c_1$	-2.421	0.159	-0.083	0.487	-0.001
$c_2$	1.599	-0.116	0.066	-0.321	0.001	$c_2$	1.151	-0.083	0.048	-0.231	0.001
$c_3$	-91.312	0.676	-0.321	18.808	-0.053	$c_3$	-65.737	0.487	-0.231	13.540	-0.038
$c_4$	0.259	-0.002	0.001	-0.053	0.00015	$c_4$	0.186	-0.001	0.001	-0.038	0.00011
<b><math>T=2.0</math></b>	$c_0$	$c_1$	$c_2$	$c_3$	$c_4$	<b><math>T=2.0</math></b>	$c_0$	$c_1$	$c_2$	$c_3$	$c_4$
$c_0$	666.811	-5.057	2.404	-137.286	0.389	$c_0$	727.929	-6.446	2.582	-149.233	0.416
$c_1$	-5.057	0.332	-0.174	1.016	-0.003	$c_1$	-6.446	0.271	-0.136	1.308	-0.004
$c_2$	2.404	-0.174	0.099	-0.483	0.001	$c_2$	2.582	-0.136	0.075	-0.524	0.001
$c_3$	-137.286	1.016	-0.483	28.277	-0.080	$c_3$	-149.233	1.308	-0.524	30.604	-0.085
$c_4$	0.389	-0.003	0.001	-0.080	0.00023	$c_4$	0.416	-0.004	0.001	-0.085	0.00024

computed for two different source-to-site distances,  $R_1=260$  km and  $R_2=480$  km; at last (3) the ratio between  $GMI_S$  estimated, which allows combining the effect of the geometrical spreading and the anelastic attenuation (Castro et al. 1990), was determined by the following equation. A sufficiently large variance was adopted for this coefficient to obtain the best value from the regression,  $\sigma' [c_4]=E' [c_4] / 1.7$ .

$$E' [c_4] = \frac{\ln(Y_1/Y_2) - E' [c_3]\ln(R_1/R_2)}{R_1 - R_2} \tag{3}$$

4. Finally, the *prior* expected value of the  $\sigma_e$  model was assigned based on the results reported in several studies (e.g., Joyner and Boore 1988; Ordaz et al. 1994; Jaimes et al. 2006), adopting  $E' [\sigma_e]=0.7$  and  $\sigma' [\sigma_e]=E' [\sigma_e] / 1.7$ .

The *posterior* expected values of  $E'' [c_i]$  and  $E'' [\sigma]$  were obtained through a Bayesian linear regression using the optimization algorithm provided by Salvatier et al. (2016).

### 4.3 Uncertainty in the regression coefficients

As mentioned earlier, the dataset used in the regression analysis presents magnitude-distance regions poorly sampled, particularly for  $M_w>7.5$  and  $R_{rup}>350$  km (see Fig. 2). This lack of data raises concerns regarding the reliability of GMPM when used to predict the GMI value for events whose parameters magnitude and distance are beyond the well-sampled region of the dataset. Therefore, in this study, it was decided to use the method proposed by Arroyo and Ordaz (2011) to quantify how the uncertainty associated with GMPM is affected due to the quality of the dataset. This method evaluated the epistemic uncertainty using the predictive variance,  $\sigma_p^2$ , which depends on the sample, on the regression technique, and on the adopted functional form. Equation (4) shows that the  $\sigma_p^2$  of a forecasted value depends on two factors: (1) the earthquake-to-earthquake variability,  $\sigma_e^2$ , contained in the database and (2) the uncertainty in the regression coefficients given by  $Z \text{COV}(\hat{c}) Z^T$ .

$$\sigma_p^2 = \sigma_e^2 + Z\text{COV}(\hat{c})Z^T \tag{4}$$

where  $\sigma_p^2$  is the total forecasted variance considering the extra uncertainty due to the quality of the dataset,  $\sigma_e^2$  the variance of the interevent residuals,  $Z$  is a  $n_p$  row vector with the GMPM parameters (see Eq. 2) for which a specific value,  $Y$ , is being forecasted and  $\text{COV}(\hat{c})$  is the covariance matrix of the coefficients given by,

$$\text{COV}(\hat{c}) = \frac{1}{n_o - n_p - 2} \left[ (X^T \Phi^{-1} X)^{-1} \left( (Y - X\hat{c})^T \Phi^{-1} (Y - X\hat{c}) \right) \right] \tag{5}$$

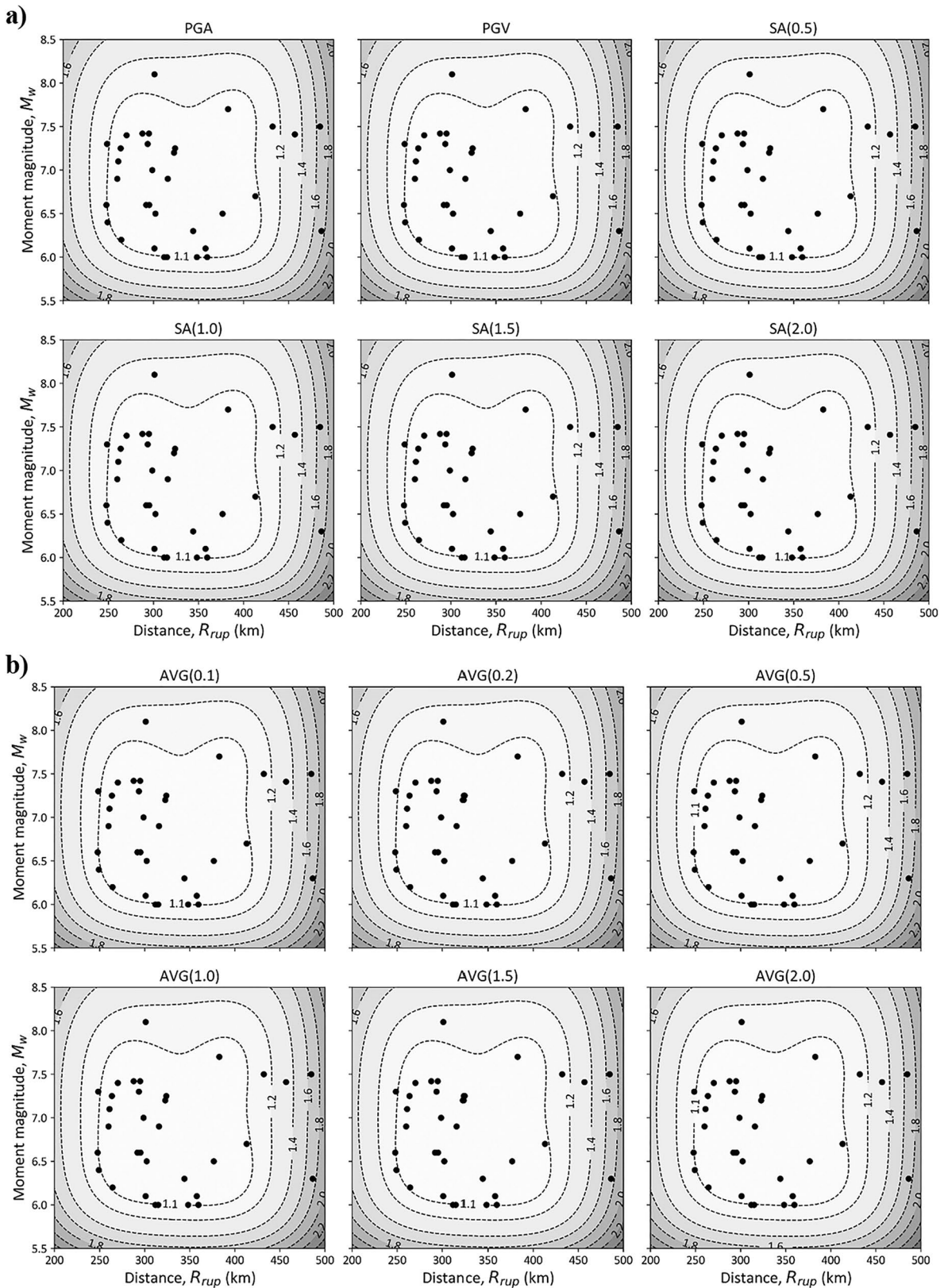
where  $X$  is a  $n_o \times n_p$  matrix with  $n_o$  observations and  $n_p$  parameters considered in the model,  $Y$  is a  $n_o \times 1$  matrix with the predicted values for the parameter given by  $X$ ,  $\hat{c}$  is a  $n_p \times 1$  matrix with the coefficients estimated from the regression analysis, and  $\Phi$ , in this case, is an identity matrix of size  $n_e$  due to only the interevent variability being analyzed (see Arroyo and Ordaz 2011). Finally,  $n_e$  is the number of earthquakes in the dataset.

Through Eq. (5) was possible to quantify the uncertainty in the regression coefficients for the GMPM derived. As will be seen later, the total forecasted variability obtained was larger for the cases where the dataset is poorly sampled and smaller enough for the opposite cases.

## 5 Results and discussion

### 5.1 Regression coefficients

Table 2 lists the attenuation coefficients and the standard deviation values obtained from the regression for PGA, PGV, SA, and AvgSA. For simplicity, Table 2 shows the coefficients for 41 and 30 structural periods for SA and AvgSA, respectively. The complete table of the regression coefficients is provided as an electronic supplement. Figures 3a and b compare the *prior* and *posterior* period-dependent coefficients estimated for SA and AvgSA, respectively. In these figures, it is noted that the *prior* and *posterior* coefficients follow a similar shape with significant differences in their amplitude, especially in the coefficients  $c_0$ ,  $c_2$ , and  $c_4$ , where the updated values are so far from those computed theoretically, which means that our *prior* knowledge of these coefficients is vague and wrong. Regarding the total standard deviation, it can be noted that the dispersion ( $\sigma$ ) calculated



◀**Fig. 4** Contours of  $s$  values, the ratio between the standard deviation of the total forecasted variability and the variability of the regression ( $s = \sigma_p/\sigma_e$ ). **a** Intensity measures PGV, PGA, and SA (0.2, 0.5, 1.0, 1.5, 2.0). **b** Intensity measures AvgSA (0.1, 0.2, 0.5, 1.0, 1.5, 2.0). The dashed lines represent,  $s$  contours, and the dots are the observed values used in the regression analysis

for AvgSA is lower than that from SA at the period range between 0.1 s and 4.0 s. Therefore, AvgSA has a better predictive capacity than SA, which has an important impact on PSHA as it contributes to a more accurate estimation of the hazard curves at small rates of exceedance.

## 5.2 Covariance matrixes and comparison of variances

The covariance matrixes needed to compute the forecasted variability, see Eq. (5), are shown in Table 3 for some IMs such as PGA, PGV, SA (0.2, 0.5, 1.0, 1.5, 2.0), and AvgSA (0.1, 0.2, 0.5, 1.0, 1.5, 2.0). As stated before, with these matrixes and the parameters to be forecasted, the epistemic uncertainty due to the sampling of the dataset could be considered in the estimation of the hazard. The supplementary material shows an extended version of Table 3 for all IMs reported in Table 2.

Moreover, to see the impact of adding the epistemic uncertainty due to the dataset sampling, a ratio between the standard deviation from the total forecasted and the one coming from the regression analysis,  $s = \sigma_p/\sigma_e$ , is computed for some intensity measures of SA and AvgSA, and it is shown in Fig. 4a, b, respectively. The dash lines correspond to the contour of  $s$  values, while the dots are the observations of the dataset (see Table 1). As it can be seen,  $s$  values close to 1 occur in the vicinity of the well-sampled region of the dataset, while higher  $s$  values (which means higher values of total forecasted variability) occur when there is a lack of data. As will be seen later, this uncertainty increase directly impacts the hazard assessment estimated for the site.

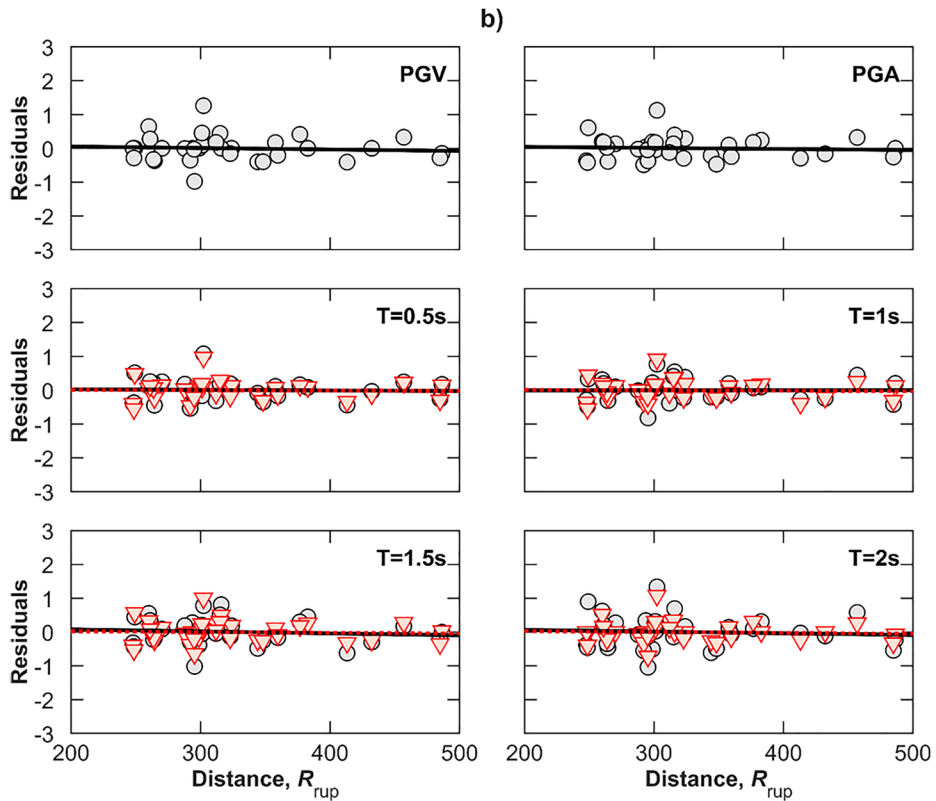
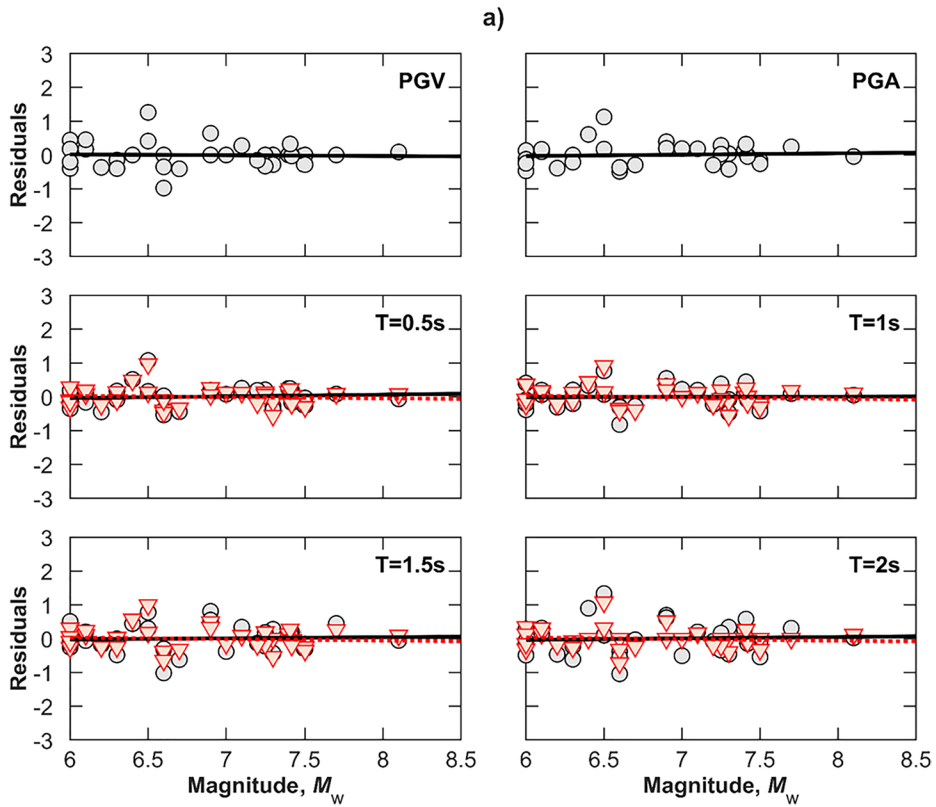
## 5.3 Residual analysis

The residuals as a function of magnitude ( $M_w$ ) and distance ( $R_{rup}$ ) are shown in Fig. 5a, b, respectively, for PGA, PGV, SA, and AvgSA. For the last two IMs, the residuals are shown for four different periods of vibration that correspond to 0.5 s, 1 s, 1.5 s, and 2 s.

The plotted residuals represent the difference between the natural logarithms of the observed data and the predicted SA (circles) and AvgSA (triangles) values. The solid and dotted lines correspond to the linear fit between the residuals computed and the  $M_w$  and  $R_{rup}$  values, respectively. As can be seen, the mean residuals obtained for  $M_w$  and  $R_{rup}$  are almost negligible for PGA, PGV, and SA, which means that the estimated coefficients of the GMPMs are adequate and provide a reasonable estimation of the GMI. Similar trends are obtained for the rest of the periods of vibration considered in this study for SA. For the case of AvgSA, the slope decreases slightly as  $M_w$  increases, especially for  $T > 1$  s. This behavior may be related to the number of events used in the regression process to obtain the coefficients for long periods since, as mentioned earlier, the number of events in the dataset is reduced by approximately 23% to estimate AvgSA at  $T \geq 1.8$  s.

## 5.4 Median predictions

Figure 6 compares the observations against the predicted values of PGV, PGA, SA, and AvgSA for two different magnitudes ( $M_w$ ) as a function of the source-to-site distance ( $R_{rup}$ ). For SA and AvgSA, the comparison is carried out for four different periods of vibration that correspond to 0.5 s, 1 s, 1.5 s, and 2 s. In this figure, the black and red solid lines represent the median prediction (50th percentile) for a magnitude value ( $M_w$ ) of 6.5 and 7.5, respectively, and the dashed lines correspond to the 16th and 84th percentiles. Likewise, the filled circles correspond to the  $6.3 \leq M_w < 6.7$  observations and the filled triangles for  $7.3 \leq M_w < 7.7$ . Significant data variability for the evaluated magnitude ranges can be observed, especially for  $6.3 \leq M_w < 6.7$ , whose dispersion increases as the period increases. This dispersion is also observed in PGV, whose intensity measure is more sensitive to the low-frequency amplitudes of the ground motion. Likewise, it is noted that the dispersion between the observations and predicted values for AvgSA is lower than the one from SA, which leads a better predictive capacity of the GMPM-AvgSA. This dispersion increase over long periods is likely related to the characteristics of the seismic recording equipment operated at the station CU since most have been short-period accelerometers with





◀**Fig. 5** Residuals in natural logarithm of PGV, PGA, SA (circles), and AvgSA (triangles) as a function of **a**  $M_w$  and **b**  $R_{rup}$ . The solid black line represents the linear fit for PGV, PGA, and SA residuals. Similarly, the dotted red line represents the linear fit for AvgSA

low sensitivity at frequencies below 1 Hz. Despite this, it is observed that the empirical data is relatively well represented by the GMPM proposed.

Figure 7 compares the response spectra observed (solid black line) with the median estimated (dotted red line) at the station CU for SA and AvgSA from the GMPM derived. Only 12 of the most intense earthquakes in the dataset are compared. As expected, the suitability of the fit varies from event to event, showing a good agreement for most of them. In the case of the events: 5 ( $M_w$  7.7), 12 ( $M_w$  6.9), 15 ( $M_w$  7.3), and 26 ( $M_w$  7.25), there is a significant difference between the observed and predicted values for both SA and AvgSA; however, the fitting is reasonable given the uncertainty of the GMPM. On the other hand, it can be noted that the spectral shape of SA is strongly influenced by the site effects present at station CU, which yield a large spectral amplification at the period range from 1 to 3 s and, consequently, widen the spectral shape of AvgSA to long periods. This condition is particularly atypical at sites classified as firm soils, where it is expected that the peak amplitude has a place at short periods below 1 s. However, in the *firm zone* of Mexico City, this amplification is caused by the rather shallow (<1 km) sediments that lie below the volcanic rocks that cover the hill zone of the Valley of Mexico (Ordaz et al. 1992).

### 5.5 Comparison of standard deviations

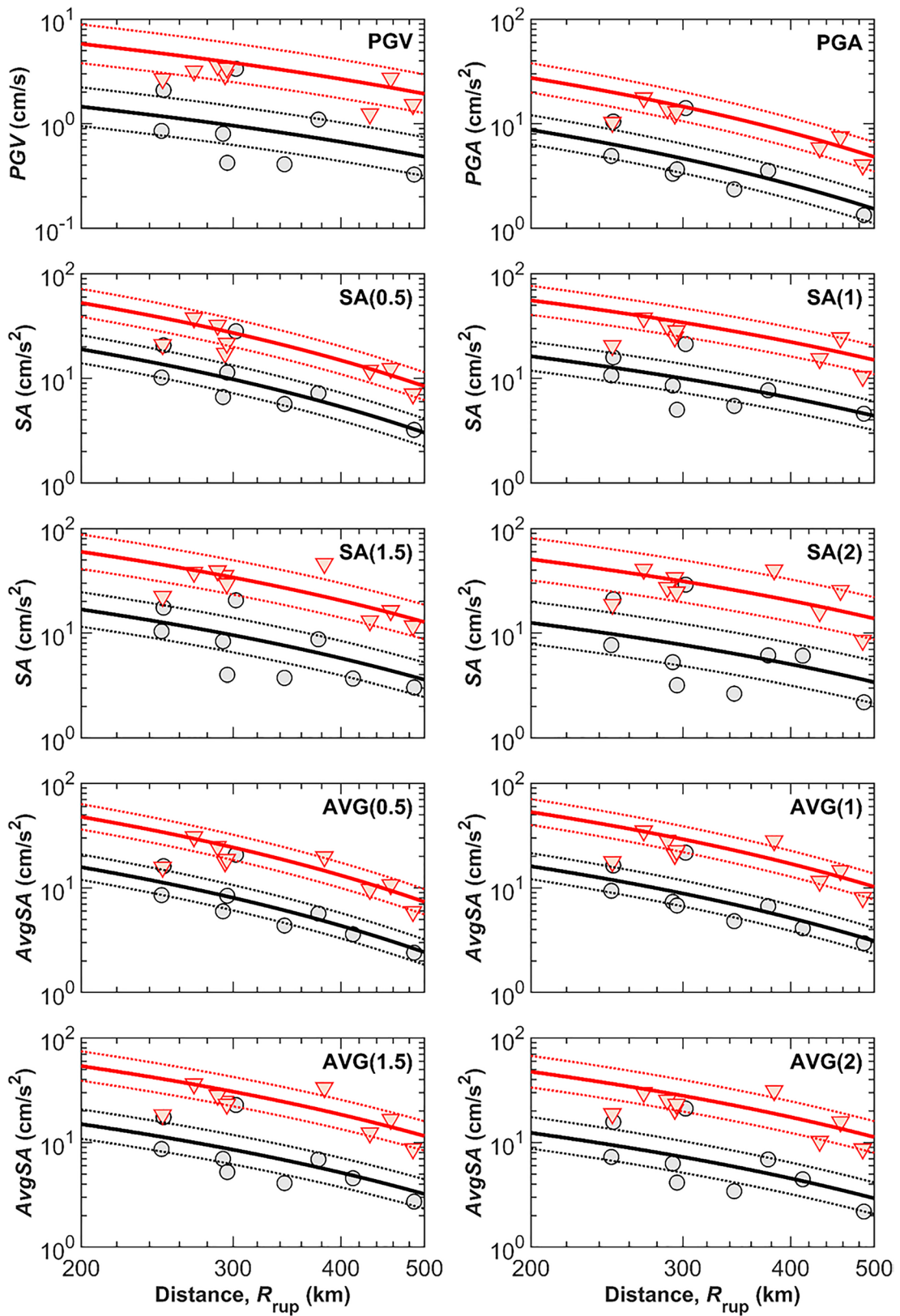
Figure 8a compares the total standard deviation of PGA and SA ( $\sigma_{\ln IM}$ ) obtained from the GMPM proposed with those obtained from two previously developed models for the station CU, corresponding to R99 and J06. This figure shows that the  $\sigma_{\ln SA}$  computed in this study is less than estimated by R99 and J06 models at all periods of vibration. Note that the PGA value is associated with  $T=0$  s. On the other hand, Fig. 8b clearly shows the percentage reduction achieved for  $\sigma_{\ln IM}$  regarding the values reported at the two GMPMs evaluated. This percentage varies from 5 to 60% for R99 and from 0.3 to 58% for J06. For both models, the lowest value occurs in the period range

from 1 to 3 s. In general, a greater reduction of the dispersion of PGA and SA is obtained for the R99 model compared to that of J06. These differences are related to the number of earthquakes used in the regression approach since CU21 contains 18 more events than R99 and 13 more than J06. These results indicate a more significant predictive capacity of the proposed GMPM than previous models.

### 5.6 Comparison with previous models

To evaluate the performance of the proposed GMPM concerning previous models developed for the station CU, Fig. 9 presents the differences in natural logarithm between the predictions of GMPM-SA proposed and the R99-SA model. Similarly, Fig. 10 shows the differences regarding the J06-SA model. In both figures, the results are presented through surface plots that allow to quickly evaluate the behavior of the residuals obtained for a wide range of magnitude values ( $6.0 \leq M_w \leq 8.5$ ) and periods of vibration ( $0 \leq T \leq 6$ ), associated with a set of four different distances. These distances correspond to 250, 300, 350, and 450 km. Likewise, a color scheme ranging from blue to red facilitates interpreting the results; the blue tones indicate that the predicted values are lower regarding the evaluated model, and the red tones indicate the opposite. In both figures, it can be observed that for  $T < 3$  s, the GMPM proposed predicts intensities below R99 and J06 for the magnitude range of 6 to 8 and distances less than 350 km. In this zone, the differences are less than 0.2 log units. For  $T > 3$  s, the model predicts intensities above R99 and J06 for the entire range of magnitude and distance. The differences are greater than 1 log unit in this area, especially at  $T > 5$  s. In general, a smaller difference is obtained regarding the J06 model than that of J06. These differences are mainly attributed to the dataset size since CU21 includes thirteen earthquakes not considered in J06 and eighteen more than R99, which provide essential information about the GMI for events with magnitudes ( $M_w$ ) between 6.0 and 7.4, and distances ( $R_{rup}$ ) between 260 and 400 km.

Another point that was evaluated of the GMPM was its performance in estimating the ground motion in other sites of the hill zone of Mexico City. For that purpose, a dataset of 24 events recorded at nine stations located southwest of Mexico City was used. These stations present a similar condition of the



◀**Fig. 6** Regression data compared with attenuation curves computed for two different magnitudes:  $M_w=6.5$  (black lines) and  $M_w=7.5$  (red lines). Solid and dashed lines represent the median and the 16th–84th percentiles. The circles and triangles filled correspond to the data observed in the range of  $6.3 \leq M_w < 6.7$  and  $7.3 \leq M_w < 7.7$ , respectively

ground motion amplification that agrees with the reported by Montalvo-Arrieta et al. (2002, 2003). Table 4 lists the stations and earthquakes employed in the analysis. In total, 105 pairs of records (two horizontal components) were obtained from these stations.

The maximum likelihood methods LH and LHH proposed by Scherbaum et al. (2004, 2009) were chosen for testing the GMPMs against the data observed. These methods allow quantifying in statistical terms the similarity between the observed and predicted ground-motion data and have been used in several studies to identify suitable models for specific seismic regions (Beauval et al. 2012; Delavaud et al. 2012). These methods were used to evaluate the GMPM performance for PGA and SA at nineteen periods between 0.1 and 6.0 s.

To apply the LH method is necessary to obtain the normalized residual for the set of observed and predicted values and compute the likelihood of an observation to be equal to or larger than a  $z$  value through the equations,

$$LH = Erf\left(\frac{|z|}{\sqrt{2}}, \infty\right) \tag{6}$$

$$z = \frac{x - \mu}{\sigma} \tag{7}$$

where  $z$  represents the normalized residual,  $x$  is the observed data, and  $\mu$  and  $\sigma$  are the mean and the standard deviation of the GMPM, respectively.  $Erf(\cdot)$  is the error function while integrating both tails of the standard normal distribution. LH can take values between 0 and 1. An LH value of 0.5 means that the observed and predicted data match their mean and standard deviation perfectly. Likewise, Scherbaum et al. (2004) provide a ranking criterion to quantify the overall model capability for predicting the observed data, assigning a category A, B, C, or D as a function of the LH obtained (refer to Scherbaum et al. (2004) for more details).

For its part, the LLH method is based on the log-likelihood approach to measuring the distance between two continuous probability density functions  $f$  and  $g$ , log-normally distributed, where  $f$  represents the distribution of the observed values and  $g$  the distribution of the GMPM evaluated. This approach computes the average log-likelihood of the considered GMPM using the observed dataset through the equation,

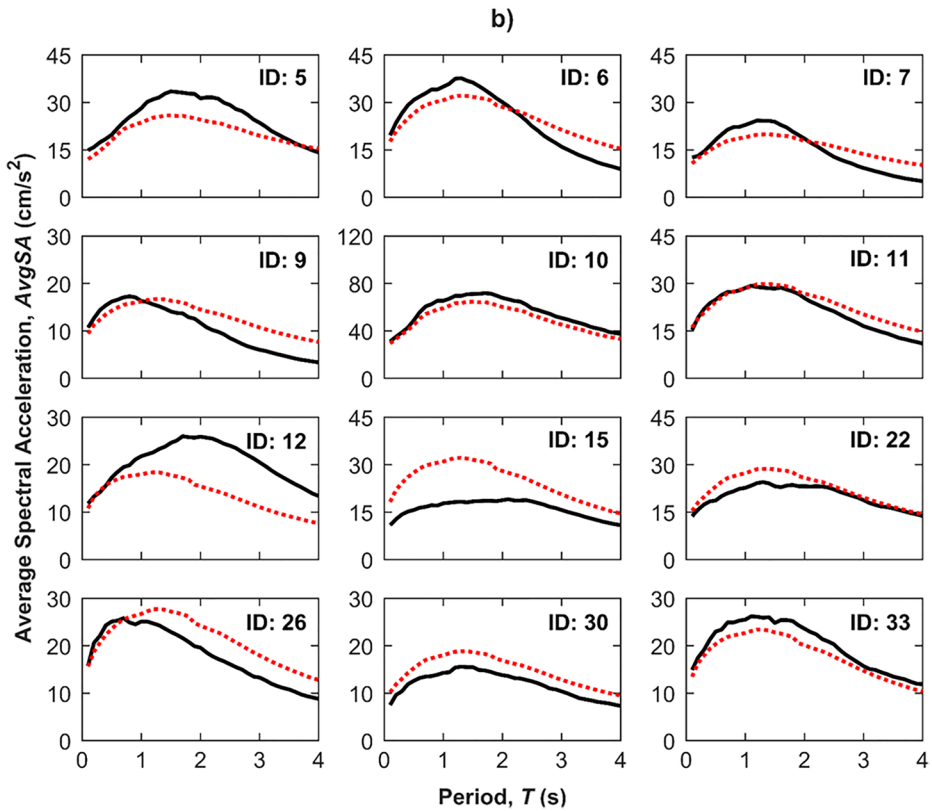
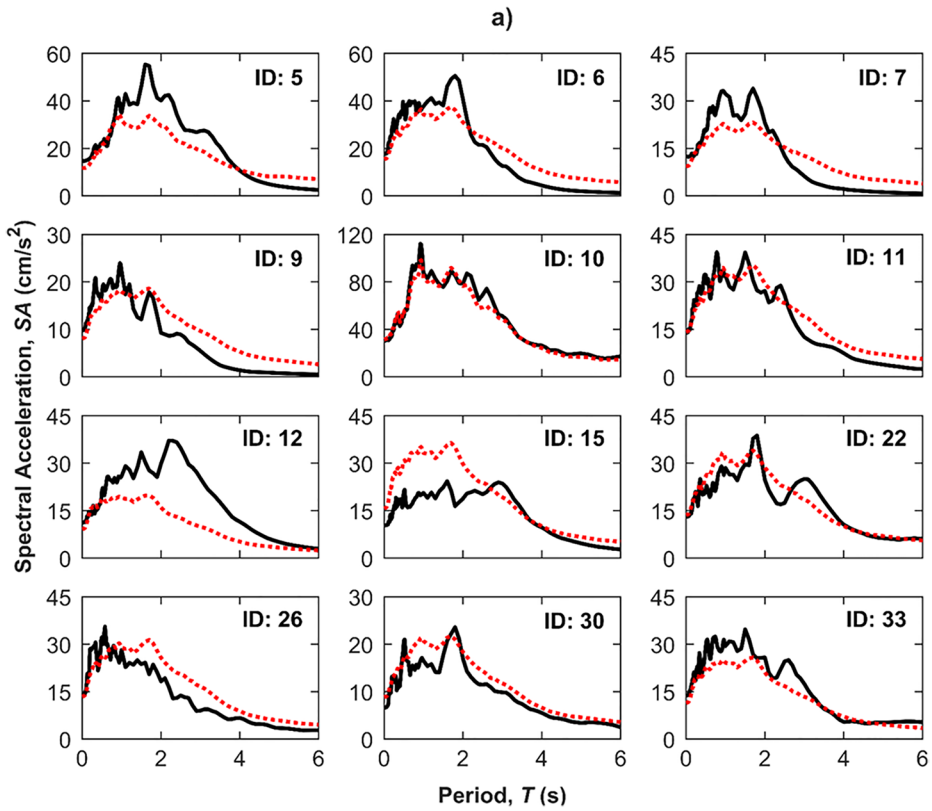
$$LLH = -\frac{1}{N} \sum_{i=1}^N \log_2(g(x_i)) \tag{8}$$

where  $N$  represents the total number of observations,  $x$  and  $g$  is the probability density function predicted by the model. A small LLH value indicates a good fit between the observed and predicted data.

Figure 11 shows the general performance of the GMPMs evaluated at the selected period range, and Table 5 lists the corresponding LH and LHH ranking indices. From this figure, the R99 GMPM reflects the best fit between the observed and predicted data, followed by J06 and the model proposed in this study. The largest differences can be observed at short periods, below 1 s, where the model derived has the lower values of LH and highest values of LLH, which means a lower capability to predict the ground motion intensity at the entire *hill zone* of Mexico City. Table 5 shows that the models proposed for the period range between 0.3 and 0.7 s have a low low predictive capacity. The difference between the GMPMs is less evident at long periods and shows a more stable ranking result.

The low predictive capacity of the proposed model, obtained from the LH and LLH methods, is likely related to the sigma of the model since its value is lower than that reported by the R99 and J06 models (see Fig. 8). Kale and Akkar (2013) point out that two GMPMs with a similar median estimation evaluated with the LH and LLH methods would benefit the one with the highest sigma. However, in terms of a PSHA, the models with larger sigma lead to a large seismic hazard at long periods.

Goodness-fit measures: median likelihood value ( $\widehat{LH}$ ), and the median, mean and standard deviation of the normalized residual ( $\widehat{NR}$ ,  $NR$ , and  $\sigma_{NR}$ ).  $\widehat{LHH}$  represent the median log-likelihood value and RK the



◀**Fig. 7** Comparison between observed and computed response spectra at the station CU for **a** SA and **b** AvgSA, using the GMPM derived. The solid black line represents the response spectrum observed, and the dotted red line the response spectrum estimated

ranking assigned. The capital letters A, B, C and D stands the predictive capability of the GMPM: high, medium, low, and unacceptable in that order.

### 5.7 Comparison of hazard curves and uniform hazard spectra

Figure 12 compares the empirical and theoretical hazard curves computed from a probabilistic seismic hazard analysis (PSHA) for PGA and SA at five structural periods that correspond to 0.2 s, 0.5 s, 1 s, 1.5 s, and 2 s. The empirical curve is obtained by counting the number of times per year a given IM value has been exceeded, dividing by the observation period that, in this case, corresponds to 57 years for PGA, SA ( $T \leq 2.7$ ), and AvgSA ( $T \leq 1.8$ ), and 37 years for PGV, SA ( $T > 2.7$ ) and AvgSA ( $T > 1.8$ ). For these IMs, the observation period decreases because only 24 events recorded from 1985 to 2021 were used to derive the GMPM. The PSHA analysis is performed just for interface earthquakes along the Mexican Pacific Coast. The characteristics of the seismogenic zones (i.e., geometry and seismicity) are taken from the current version of Mexico's Seismic Design Code of the Federal Electricity Commission (MDOC CFE 2015). The attenuation models used correspond to R99, J06, and the one proposed in this study. The computations were made using R-CRISIS (Ordaz et al. 2021), based on the classic Esteva–Cornell approach (McGuire 2008).

The empirical curves were computed for 30 intensity values logarithmically separated by equal intervals between the limits 1 to 1000  $\text{cm/s}^2$  for the SA curve (Fig. 11) and 1 to 100  $\text{cm/s}^2$  for the AvgSA curve (Fig. 12). Therefore, their ordinates are the same in the six figures associated with a specific IM. In these figures, most of the empirical hazard curves, compared with the theoretical hazard curves computed from PSHA, tend to saturate at intensities below 5  $\text{cm/s}^2$ , which may be related to the fact that earthquakes with low ground-motion intensities were dismissed from the dataset. The same behavior occurs

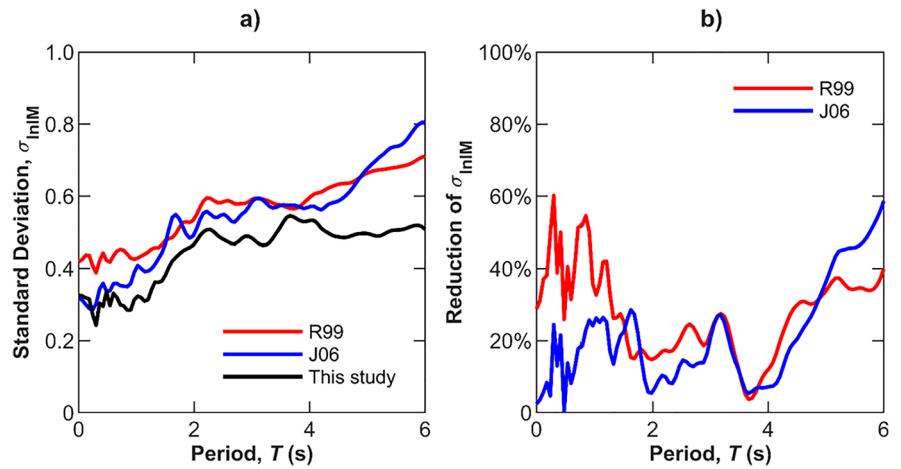
at intensities above 30  $\text{cm/s}^2$  because there is only one event in the dataset that generates intensities above this threshold, so its exceedance rate corresponds to the maximum observation period, equivalent to  $v(a) = 0.0175$  for PGA, SA ( $T \leq 2.7$ ) and AvgSA ( $T \leq 1.8$ ), and  $v(a) = 0.027$  for PGV, SA ( $T > 2.7$ ) and AvgSA ( $T > 1.8$ ). For moderate intensities, the differences between both hazard curves are not so large and seem fit for the observed data. Similar observations were made by Ordaz and Reyes (1999).

Figure 12 shows that the shape of the hazard curves estimated with the proposed model disregarding the uncertainty in the regression coefficients (herein, HDU, dotted black line) follows a similar shape to the estimated with the R99 (continuous gray line) and J06 (dotted red line) models. The three GMPMs compute similar exceedance rates,  $v(a)$ , for low intensities. However, their differences increase as the intensity increases. Compared to the proposed model, the R99 model estimates higher exceedance rates for PGA and SA for all periods evaluated, except for  $T = 2$  s. On the other hand, the J06 model estimates lower exceedance rates for PGA and SA at  $T = 0.2$  s and  $T = 0.5$  s, the opposite case occurs for  $T = 1.0$  s and  $T = 1.5$  s. Like R99, the forecasts of J06 at  $T = 2$  s are the same as those of the proposed model.

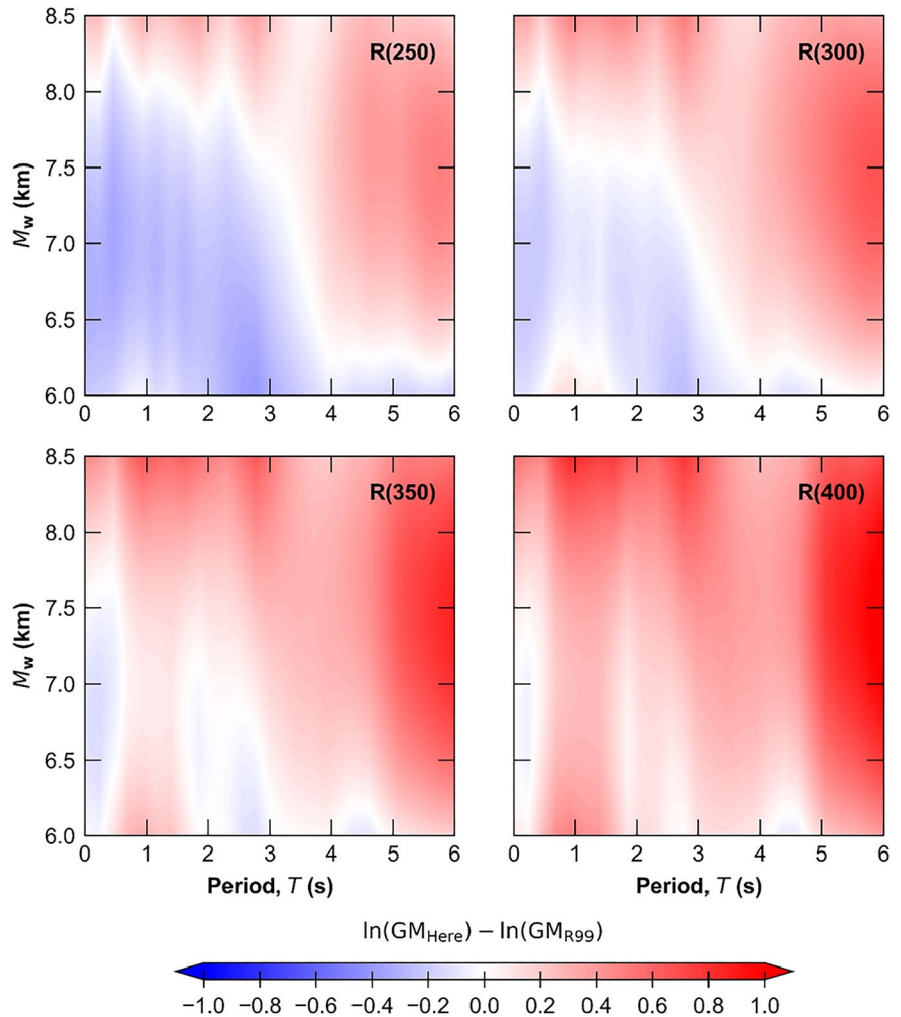
Likewise, in Fig. 12, the hazard curves are also compared with the proposed model, considering the uncertainty in the regression coefficients (herein, HCU, continuous black line). As expected, for small and large intensities, the hazard level of HDU is larger than HCU. This increment at small intensities, with  $v(a) \geq 0.1$ , is produced by the additional epistemic uncertainty for small magnitude events ( $M_w < 6$ ) at short distances ( $R_{rup} < 240$  km) whose predictive standard deviation,  $\sigma_p$ , is between 20 and 100% more than the obtain directly from the Bayesian regression,  $\sigma_e$  (see Fig. 4). The same behavior occurs for large intensities, with  $v(a) \leq 0.01$ , since the hazard level is controlled by large events with  $M_w > 7.5$  and  $R_{rup} > 280$  km, in which  $\sigma_p$  is between 10 and 80% more than  $\sigma_e$ . The differences become less evident at moderate intensities, with  $0.1 \geq v(a) \geq 0.01$ , whose hazard level is controlled by earthquakes with magnitudes ( $M_w$ ) between 6.0 and 7.5 and distances ( $R_{rup}$ ) between 250 and 350 km. These parameters fall in the range where  $\sigma_e^2$  and  $\sigma_p^2$  are similar (see Fig. 4) and correspond to the well-sampled region of the dataset. The same effect is observed in the hazard curves



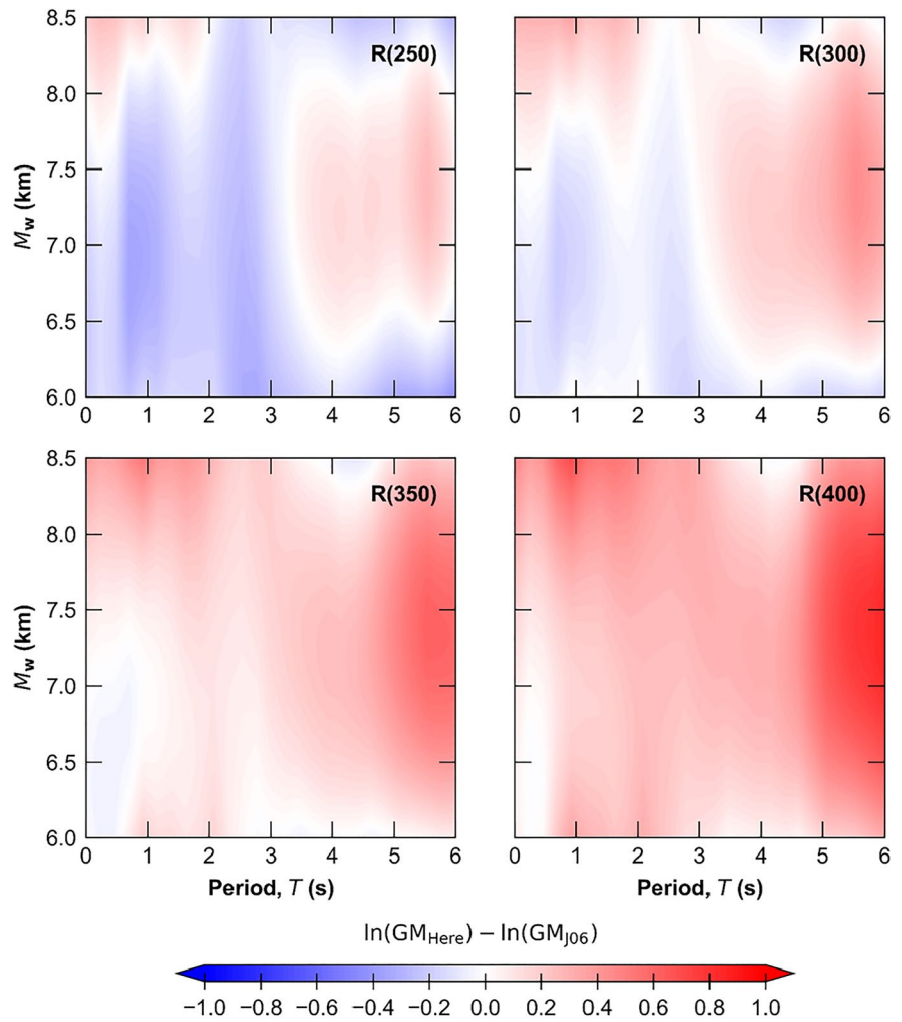
**Fig. 8** Comparison between standard deviation values from R99, J06, and the GMPM proposed for PGA and SA



**Fig. 9** Differences in terms of natural logarithm between predictions of GMPM-SA proposed and R99-SA model. The differences are presented for a set of four distances ( $R_{rup}$ ) as a function of the period of vibration ( $T$ ) and magnitude ( $M_w$ )



**Fig. 10** Differences in terms of natural logarithm between predictions of GMPM-SA proposed and J06-SA model. The differences are presented for a set of four distances ( $R_{rup}$ ) as a function of the period of vibration ( $T$ ) and magnitude ( $M_w$ )

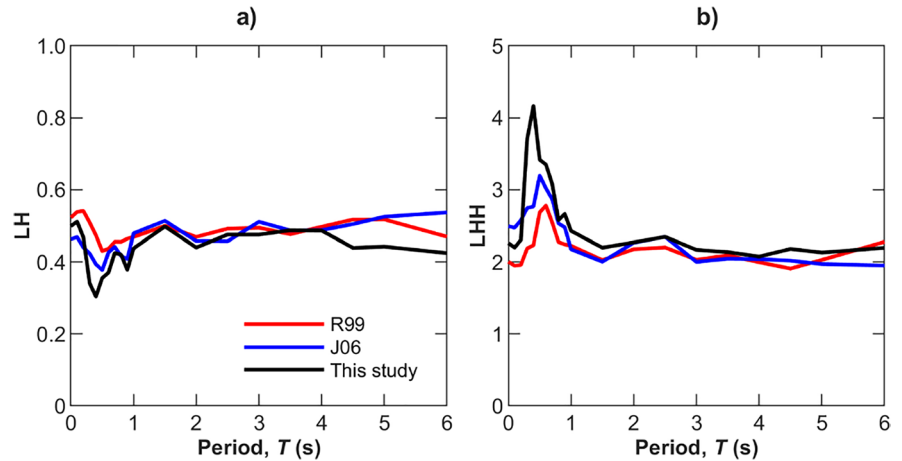


**Table 4** Stations and earthquakes used in the analysis

Station	Latitude (°N)	Longitude (°W)	Altitude (m)	Event recorded*
TACY	19.4045	99.1952	2240	10, 11, 12, 15, 16, 18, 19, 20, 22, 29, 30, 32, 33
PA34	19.2016	99.0491	2240	12, 14, 15, 16, 18, 20, 22, 26, 27, 30, 31, 32, 33
IM40	19.3428	99.2032	2365	16, 18, 19, 22, 23, 30
MT50	19.4253	99.1900	2234	12, 14, 17, 18, 19, 20, 21, 22, 23, 24, 26, 27, 20, 21, 22, 23, 24
UI21	19.3700	99.2642	2540	15, 18, 21, 22, 23, 24, 26, 27, 28, 21, 22, 23, 24
FJ74	19.2990	99.2100	2240	12, 13, 14, 15, 16, 18, 22, 24, 25, 27, 28, 32, 33,
TE07	19.4269	99.2217	2290	15, 18
TP13	19.2922	99.1708	2265	12, 15, 16, 18, 19, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33
CS78	19.3656	99.2262	2430	14, 15, 19, 21, 22, 25, 30, 31, 32, 33

\*The seismic source parameters of the events are listed in Table 1

**Fig. 11** Performance of the GMPMs evaluated at the selected period range. **a** LH values vs period and **b** LLH values vs period



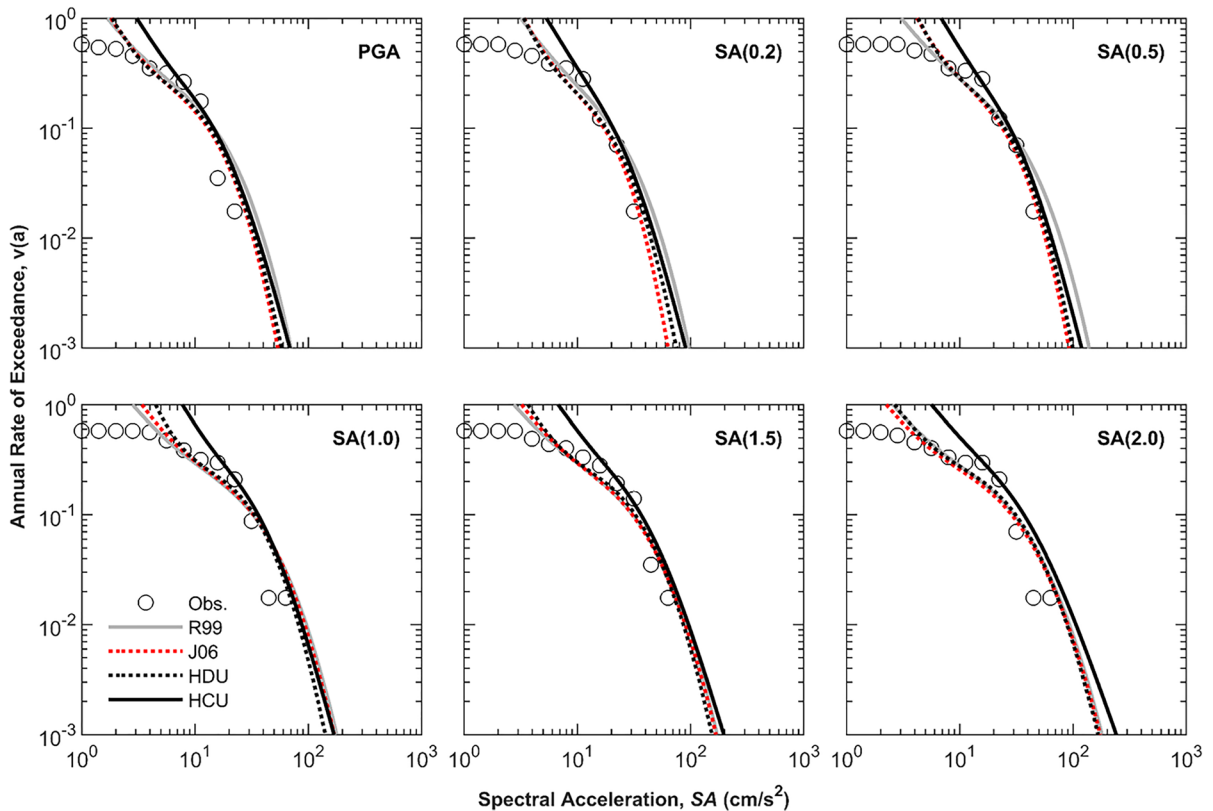
computed for AvgSA (see Fig. 12), although less evident at low intensities than in SA curves.

Finally, Fig. 13 compares the empirical and theoretical hazard curve computed from PSHA for AvgSA at six periods corresponding to 0.1 s, 0.2 s, 0.5 s, 1 s, 1.5 s, and 2 s. In this case, the hazard curves are

estimated directly from GMPM-AvgSA proposed, disregarding the uncertainty in the regression coefficient (HDU), and compared with those obtained indirectly using the existing GMPM-SAs (i.e., R99 and J06) and a SA inter-period correlation model proposed by Rodríguez-Castellanos et al. (2021). The

**Table 5** Ranking of the GMPM evaluated

IMT	Reyes (1999)						Jaimes et al. (2006)						This study					
	RK	LH	NR	NR	$\sigma_{NR}$	LHH	RK	LH	NR	NR	$\sigma_{NR}$	LHH	RK	LH	NR	NR	$\sigma_{NR}$	LHH
PGA	A	0.52	-0.18	-0.06	0.97	2.00	C	0.46	0.10	0.17	1.27	2.49	A	0.50	0.08	0.23	1.12	2.26
0.1	A	0.54	-0.19	-0.12	0.92	1.95	C	0.47	-0.06	0.03	1.27	2.47	A	0.51	0.07	0.21	1.08	2.20
0.2	A	0.54	-0.21	-0.19	0.92	1.96	C	0.44	0.07	0.06	1.32	2.58	B	0.47	0.15	0.16	1.16	2.30
0.3	A	0.51	0.14	0.23	1.07	2.19	C	0.42	0.44	0.51	1.31	2.75	D	0.34	0.79	0.90	1.58	3.70
0.4	B	0.47	0.48	0.46	1.03	2.23	D	0.39	0.67	0.79	1.18	2.77	D	0.30	1.03	1.26	1.54	4.17
0.5	C	0.43	0.57	0.68	1.20	2.69	D	0.38	0.61	0.84	1.38	3.20	D	0.36	0.82	0.93	1.43	3.42
0.6	D	0.44	0.69	0.81	1.17	2.78	C	0.43	0.58	0.71	1.36	3.02	D	0.37	0.84	0.94	1.39	3.35
0.7	C	0.46	0.62	0.69	1.11	2.54	C	0.44	0.33	0.54	1.36	2.87	D	0.42	0.64	0.81	1.34	3.08
0.8	C	0.46	0.57	0.53	1.02	2.28	B	0.42	0.39	0.34	1.25	2.53	B	0.42	0.43	0.46	1.24	2.58
0.9	B	0.46	0.37	0.43	1.05	2.24	B	0.41	0.13	0.21	1.25	2.48	C	0.38	0.31	0.32	1.33	2.66
1.0	B	0.47	0.21	0.27	1.09	2.22	A	0.48	0.08	0.19	1.07	2.18	A	0.44	0.17	0.15	1.23	2.43
1.5	B	0.50	-0.27	-0.22	0.97	2.03	A	0.51	-0.22	-0.23	0.95	2.00	B	0.50	-0.34	-0.36	1.04	2.20
2.0	C	0.47	-0.50	-0.41	1.01	2.18	B	0.46	-0.33	-0.27	1.11	2.27	C	0.44	-0.54	-0.42	1.07	2.27
2.5	A	0.49	-0.13	-0.16	1.09	2.20	B	0.46	-0.13	-0.09	1.19	2.35	B	0.48	-0.13	-0.11	1.19	2.35
3.0	A	0.49	0.16	0.10	0.99	2.03	A	0.51	0.08	0.05	0.97	2.00	A	0.48	-0.09	-0.02	1.08	2.17
3.5	A	0.48	0.06	0.09	1.03	2.09	A	0.49	0.03	0.02	1.00	2.05	A	0.49	-0.21	-0.17	1.05	2.14
4.0	A	0.50	0.18	0.20	0.95	2.00	A	0.49	0.11	0.15	0.99	2.04	A	0.49	-0.18	-0.12	1.02	2.08
4.5	A	0.52	0.19	0.19	0.88	1.91	A	0.50	0.02	0.17	0.97	2.02	B	0.44	-0.36	-0.13	1.08	2.18
5.0	B	0.52	0.30	0.39	0.91	2.03	A	0.53	0.03	0.19	0.93	1.97	B	0.44	-0.40	-0.22	1.04	2.13
6.0	C	0.47	0.58	0.63	0.96	2.28	A	0.54	0.15	0.27	0.90	1.95	B	0.42	-0.31	-0.20	1.08	2.20



**Fig. 12** Comparison between empirical and computed hazard curves for  $SA(T)$ . The circles denote the empirical curve obtained from observed data. The continuous black line represents the hazard curve computed considering the uncertainty in the regression coefficients (HCU).  $T$  and the dotted black line correspond to the hazard curve computed, disregard-

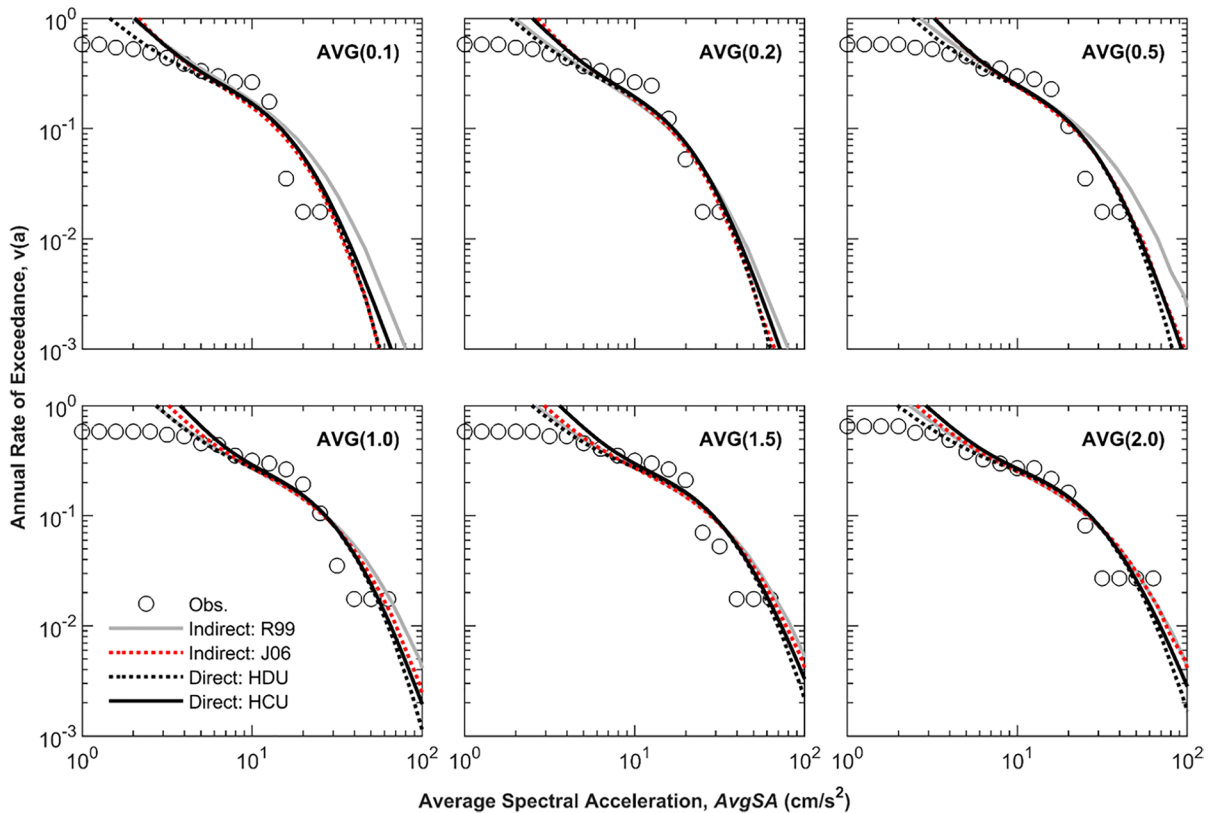
ing this uncertainty (HDU). Both curves were obtained with the GMPM proposed in this study. Likewise, the dotted red line and the continuous gray line represent the hazard curve obtained with the GMPM proposed by J06 and R99, respectively

stepped shape of the empirical curve is due to the lack of events in the database that generate large intensities. Therefore, the exceedance rates computed for large accelerations are less reliable for comparison purposes. Despite this, it is noted that the shapes of the calculated hazard curves match reasonably well with the observed data at intensities below  $20 \text{ cm/s}^2$ . The three GMPMs compute similar exceedance rates,  $v(a)$ , for low intensities. However, as observed for SA, their differences increase as intensity increases. Additionally, the better predictive capacity of AvgSA with respect to SA has an important impact on PSHA since it contributes to a more precise estimation of the hazard curve at low rates of exceedances. The R99 model estimates higher exceedance rates for all periods than the proposed mode. Likewise, the J06 model also estimates higher exceedance rates for most periods,

except for  $T=0.1 \text{ s}$  and  $T=0.2 \text{ s}$ . There are no significant differences in the seismic hazard estimations computed from the proposed model compared to previous models. However, according to the available data, the suggested GMPM estimates the seismic hazard more adequately than the other models.

### 6 Conclusions

This article presents an updated GMPM to estimate the peak ground acceleration (PGA), peak ground velocity (PGV), 5% damped pseudo-spectral acceleration (SA), and the average spectral acceleration (AvgSA) at the hill zone of Mexico City for interface earthquakes that occur along the Pacific coast of Mexico. The model is built as a function of the



**Fig. 13** Comparison between empirical and computed hazard curves for  $\text{AvgSA}(T)$ . The dotted denotes the empirical curve obtained from observed data. The continuous black line represents the hazard curve computed considering the uncertainty in the regression coefficients (HCU), and the dotted black line corresponds to the hazard curve computed disregarding this

uncertainty (HDU). Both curves were obtained directly with the GMPM proposed in this study. Likewise, the dotted red line and the continuous gray line represent the hazard curve obtained with the GMPM proposed by J06 and R99, respectively

magnitude ( $M_w$ ) and the closest distance to the rupture plane ( $R_{rup}$ ), using thirty-three reverse-faulting events recorded at the station CU from 1965 to 2021.

The new strong-motion dataset (CU21) includes thirteen more earthquakes than the one used in J06 and eighteen more than in R99, providing valuable information on the GMI for events with a magnitude range ( $M_w$ ) between 6.0 and 7.4 and a distance ( $R_{rup}$ ) between 260 and 400 km. Furthermore, the results showed that the dispersion of the GMPM proposed is lower than the previous models for PGA and SA, which means better predictability and more reliable estimates of the seismic hazard at the CU site. However, when quantifying its predictive capability to estimate the ground motion at the entire hill zone of Mexico City, the lowest ranking is obtained compared to the other tested models. It is thought that the low

predictive capacity of the proposed model, obtained from the LH and LLH methods, is related to the standard deviation ( $\sigma_e$ ) of the model since its value is lower than that reported by the R99 and J06 models. The standard deviation obtained for GMPM-AvgSA shows that the dispersion is lower than that from GMPM-SA at the period range between 0.1 s and 4.0 s, except in the range of 0.22–0.38 s where GMPM-SA presents a lower dispersion than the GMPM-AvgSA. Despite this, it is thought that the GMPM-AvgSA model has a better predictive capacity.

The analysis carried out of the uncertainties due to the sampling of the dataset showed that there is a variability that is not always taken into account in the GMPMs, and then an underestimation of the hazard levels is presented for lower and higher annual rates. This behavior explains that the source model used to



estimate the hazard curves considers events below or above the magnitude range in the dataset used to obtain the regression coefficients. Therefore, it is important to highlight these uncertainties and include them in the seismic hazard assessment.

**Acknowledgements** The authors thank the Seismic Instrumentation Unit of the Institute of Engineering of the National Autonomous University of Mexico (UNAM) for sharing the strong motion database.

**Author contribution** All authors contributed to the preparation of this paper.

**Funding** The second author received funding from the EU Horizon 2020 program under Grant Agreement Number 813137, ITN-MSCA URBASIS project.

**Data availability** The catalog of events was downloaded from Sawires et al. (2019). The data source of the finite-fault geometry information was obtained from Mai and Thingbaijam (2014) and National Seismological Service (SSN 2022). The records used were provided by the strong-motion network of the Institute of Engineering of the National Autonomous University of Mexico (UNAM), <http://aplicaciones.iingen.unam.mx/AcelerogramasRSM/>.

## Declarations

**Competing interests** The authors declare no competing interests.

**Ethics approval and consent to participate** Not applicable.

**Consent for publication** All authors give their consent so this paper can be published.

**Conflict of interest** The authors declare no competing interests.

## References

- Abrahamson NA, Silva WJ (1997) Empirical response spectral attenuation relations for shallow crustal earthquakes. *Seismol Res Lett* 68:94–109. <https://doi.org/10.1785/gssrl.68.1.94>
- Alonso L, Espinosa JM, Mora I et al (1979) Informe preliminar sobre el sismo del 14 de marzo de 1979 cerca de la costa de Guerrero. Parte A. Mexico City, Mexico
- Ancheta TD, Darragh RB, Stewart JP et al (2014) NGA-West2 database. *Earthq Spectra* 30:989–1005. <https://doi.org/10.1193/070913EQS197M>
- Arroyo D, Ordaz M (2011) On the forecasting of ground-motion parameters for probabilistic seismic hazard analysis. *Earthq Spectra* 27:1–21. <https://doi.org/10.1193/1.3525379>
- Astiz L, Kanamori H, Eissler H (1987) Source characteristic of earthquakes in the Michoacan seismic gap in Mexico. *Bull Seismol Soc Am* 77:1326–1346
- Atkinson GM (2006) Single-station sigma. *Bull Seismol Soc Am* 96:446–455. <https://doi.org/10.1785/0120050137>
- Beauval C, Tasan H, Laurendeau A et al (2012) On the testing of ground-motion prediction equations against small-magnitude data. *Bull Seismol Soc Am* 102:1994–2007. <https://doi.org/10.1785/0120110271>
- Bindi D, Massa M, Luzi L et al (2014) Pan-European ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods up to 3.0 s using the RESORCE dataset. *Bull Earthq Eng* 12:391–430. <https://doi.org/10.1007/s10518-013-9525-5>
- Castro R, Singh SK, Mena E (1988) Mexico earthquake of September 19, 1985—an empirical model to predict Fourier amplitude spectra of horizontal ground motion. *Earthq Spectra* 4:675–685. <https://doi.org/10.1193/1.1585497>
- Castro RR, Anderson JG, Singh SK (1990) Site response, attenuation and source spectra of S waves along the Guerrero, Mexico, subduction zone. *Bull Seismol Soc Am* 80:1481–1503
- Delavaud E, Scherbaum F, Kuehn N, Allen T (2012) Testing the global applicability of ground-motion prediction equations for active shallow crustal regions. *Bull Seismol Soc Am* 102:707–721. <https://doi.org/10.1785/0120110113>
- Díaz-Rodríguez JA (1989) Behavior of Mexico City clay subjected to undrained repeated loading. *Can Geotech J* 26:159–162. <https://doi.org/10.1139/t89-016>
- Díaz-Rodríguez JA (1992) On dynamic properties of Mexico City clay for wide strain range. In: Tenth World Conference on Earthquake Engineering. Madrid, Spain, pp 1257–1262
- Díaz-Rodríguez JA, Santamarina JC (2001) Mexico city soil behavior at different strains: observations and physical interpretation. *Journal of Geotechnical and Geoenvironmental Engineering* 127:783–789. [https://doi.org/10.1061/\(ASCE\)1090-0241\(2001\)127:9\(783\)](https://doi.org/10.1061/(ASCE)1090-0241(2001)127:9(783))
- Eads L, Miranda E, Lignos DG (2015) Average spectral acceleration as an intensity measure for collapse risk assessment. *Earthq Eng Struct Dyn* 44:2057–2073. <https://doi.org/10.1002/eqe.2575>
- Flores Estrella H, Aguirre González J (2003) SPAC: An alternative method to estimate earthquake site effects in Mexico City. *Geofis Int* 42:227–236. <https://doi.org/10.22201/igeof.00167169p.2003.42.2.267>
- Giovenale P, Cornell CA, Esteva L (2004) Comparing the adequacy of alternative ground motion intensity measures for the estimation of structural responses. *Earthq Eng Struct Dyn* 33:951–979. <https://doi.org/10.1002/eqe.386>
- Iglesias A, Singh SK, Ordaz M et al (2007) The seismic alert system for Mexico City: an evaluation of its performance and a strategy for its improvement. *Bull Seismol Soc Am* 97:1718–1729. <https://doi.org/10.1785/0120050202>
- Jaimes MA, Reinoso E, Ordaz M (2006) Comparison of methods to predict response spectra at instrumented sites given the magnitude and distance of an earthquake. *J*

- Earthquake Eng 10:887–902. <https://doi.org/10.1080/13632460609350622>
- Jaimes MA, Ramirez-Gaytan A, Reinoso E (2015) Ground-motion prediction model from intermediate-depth intraslab earthquakes at the hill and lakebed zones of Mexico City. *J Earthquake Eng* 19:1260–1278. <https://doi.org/10.1080/13632469.2015.1025926>
- Joyner WB, Boore DM (1988) Measurement, characterization, and prediction of strong ground motion. In: Proc Conf on Earthquake Engineering and Soil Dynamics. GT Div/ASCE, Park City, UTAH
- Kagawa T (1996) Estimation of velocity structures beneath Mexico City using microtremor array data. In: 11th World Conference on Earthquake Engineering. Acapulco, Mexico, p 8
- Kale O, Akkar S (2013) A new procedure for selecting and ranking ground-motion prediction equations (GMPEs): the Euclidean Distance-Based Ranking (EDR) method. *Bull Seismol Soc Am* 103:1069–1084. <https://doi.org/10.1785/0120120134>
- Kohrangi M, Bazzurro P, Vamvatsikos D (2016a) Vector and scalar IMs in structural response estimation, Part I: Hazard analysis. *Earthq Spectra* 32:1507–1524. <https://doi.org/10.1193/053115EQS080M>
- Kohrangi M, Bazzurro P, Vamvatsikos D (2016b) Vector and scalar IMs in structural response estimation, Part II: Building demand assessment. *Earthq Spectra* 32:1525–1543. <https://doi.org/10.1193/053115EQS081M>
- Kohrangi M, Bazzurro P, Vamvatsikos D, Spillatura A (2017) Conditional spectrum-based ground motion record selection using average spectral acceleration. *Earthq Eng Struct Dyn* 46:1667–1685. <https://doi.org/10.1002/eqe.2876>
- Kohrangi M, Kotha SR, Bazzurro P (2018) Ground-motion models for average spectral acceleration in a period range: direct and indirect methods. *Bull Earthq Eng* 16:45–65. <https://doi.org/10.1007/s10518-017-0216-5>
- Lermo J, Chávez-García FJ (1993) Site effect evaluation using spectral ratios with only one station. *Bull Seismol Soc Am* 83:1574–1594
- Liu L, Pezeshk S (1999) An improvement on the estimation of pseudoresponse spectral velocity using RVT method. *Bull Seismol Soc Am* 89:1384–1389
- López-Castañeda AS, Reinoso E (2021) Strong-motion duration predictive models from subduction interface earthquakes recorded in the hill zone of the Valley of Mexico. *Soil Dyn Earthq Eng* 144. <https://doi.org/10.1016/j.soildyn.2021.106676>
- Luco N, Cornell CA (2007) Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. *Earthq Spectra* 23:357–392. <https://doi.org/10.1193/1.2723158>
- Mai PM, Thingbaijam KKS (2014) SRCMOD: an online database of finite-fault rupture models. *Seismol Res Lett* 85:1348–1357. <https://doi.org/10.1785/0220140077>
- Mayoral JM, Romo MP, Osorio L (2008) Seismic parameters characterization at Texcoco lake, Mexico. *Soil Dyn Earthq Eng* 28:507–521. <https://doi.org/10.1016/j.soildyn.2007.08.004>
- Mayoral JM, Castañon E, Alcantara L, Tepalcapa S (2016) Seismic response characterization of high plasticity clays. *Soil Dyn Earthq Eng* 84:174–189. <https://doi.org/10.1016/j.soildyn.2016.02.012>
- Mayoral JM, Asimaki D, Tepalcapa S et al (2019) Site effects in Mexico City basin: past and present. *Soil Dyn Earthq Eng* 121:369–382. <https://doi.org/10.1016/j.soildyn.2019.02.028>
- McGuire RK (2008) Probabilistic seismic hazard analysis: early history. *Earthq Eng Struct Dyn* 37:329–338. <https://doi.org/10.1002/eqe.765>
- MDOC CFE (2015) Manual de Diseño de Obras Civiles de la CFE, Instituto de Investigaciones Eléctricas, Comisión Federal de Electricidad, Mexico (in spanish)
- Chiou BS, Youngs RR (2008) An NGA model for the average horizontal component of peak ground motion and response spectra. 24:173–215. <https://doi.org/10.1193/1.2894832>
- Meli R, Miranda E (1985) The Effect of the September 1985 Earthquakes on the Constructed Facilities of Mexico City. Part I: Structural Aspects. Mexico City, Mexico
- Montalvo-Arrieta JC, Sánchez-Sesma FJ, Reinoso E (2002) A virtual reference site for the Valley of Mexico. *Bull Seismol Soc Am* 92:1847–1854. <https://doi.org/10.1785/0120010257>
- Montalvo-Arrieta JC, Reinoso-Angulo E, Sánchez-Sesma FJ (2003) Observations of strong ground motion at hill sites in Mexico City from recent earthquakes. *Geofis Int* 42:205–217. <https://doi.org/10.22201/igeof.00167169p.2003.42.2.265>
- NTC-DC (2017) Normas Técnicas Complementarias. Gac Of Ciudad México
- Ordaz M, Reyes C (1999) Earthquake hazard in Mexico City: observations versus computations. *Bull Seismol Soc Am* 89:1379–1383. <https://doi.org/10.1785/bssa0890051379>
- Ordaz M, Singh SK (1992) Source spectra and spectral attenuation of seismic waves from Mexican earthquakes, and evidence of amplification in the hill zone of Mexico City. *Bull Seismol Soc Am* 82:24–43
- Ordaz M, Singh SK, Reinoso E et al (1988) The Mexico earthquake of September 19, 1985-estimation of response spectra in the lake bed zone of the Valley of Mexico. *Earthq Spectra* 4:815–833
- Ordaz M, Singh SK, Arciniega A (1994) Bayesian attenuation regressions: an application to Mexico City. *Geophys J Int* 117:335–344. <https://doi.org/10.1111/j.1365-246X.1994.tb03936.x>
- Ordaz M, Reinoso E, Jaimes MA et al (2017) High-resolution early earthquake damage assessment system for Mexico City based on a single-station. *Geofísica internacional* 56:117–135. <https://doi.org/10.19155/geofint.2017.056.1.9>
- Ordaz M, Salgado-Gálvez MA, Giraldo S (2021) R-CRISIS: 35 years of continuous developments and improvements for probabilistic seismic hazard analysis. *Bull Earthq Eng* 19:2797–2816. <https://doi.org/10.1007/s10518-021-01098-w>
- Ordaz M (2017) Normas de Diseño por Sismo en México DF: algunas novedades interesantes. *Alternativas* 17:106–115. <https://doi.org/10.23878/alternativas.v17i3.220>
- Orozco V, Reinoso E (2007) Revisión a 50 años de los daños ocasionados en la Ciudad de México por el sismo del 28 de Julio de 1957 con ayuda de investigaciones recientes y sistemas de información geográfica. *Revista De Ingeniería Sísmica* 87:61–87

- Ramírez-Gaytán A, Aguirre J, Jaimes MA, Huérfano V (2014) Scaling relationships of source parameters of Mw 6.9–8.1 earthquakes in the cocos-rivera-north american subduction zone. *Bull Seismol Soc Am* 104:840–854. <https://doi.org/10.1785/0120130041>
- Reinoso E, Ordaz M (1999) Spectral ratios for Mexico City from free-field recordings. *Earthq Spectra* 15:273–295
- Reyes C (1999) El estado limite de servicio en el diseño sísmico de los edificios. Phd Thesis, UNAM
- Rodríguez-Castellanos A, Ruiz SE, Bojórquez E, Reyes-Salazar A (2021) Influence of spectral acceleration correlation models on conditional mean spectra and probabilistic seismic hazard analysis. *Earthq Eng Struct Dyn* 50:309–328. <https://doi.org/10.1002/eqe.3331>
- Romo M, Ovando-Shelley E (1996) Modelling the dynamic behaviour of Mexican clays. In: XI International Conference on Earthquake Engineering. Acapulco, Mexico
- Salvatier J, Wiecki T V., Fonnesbeck C (2016) Probabilistic programming in Python using PyMC3. *PeerJ Comput Sci* 2:e55. <https://doi.org/10.7717/peerj-cs.55>
- Sawires R, Santoyo MA, Peláez JA, Corona Fernández RD (2019) An updated and unified earthquake catalog from 1787 to 2018 for seismic hazard assessment studies in Mexico. *Sci Data* 6:1–14. <https://doi.org/10.1038/s41597-019-0234-z>
- Scherbaum F, Cotton F, Smit P (2004) On the use of response spectral-reference data for the selection and ranking of ground-motion models for seismic-hazard analysis in regions of moderate seismicity: the case of rock motion. *Bull Seismol Soc Am* 94:2164–2185. <https://doi.org/10.1785/0120030147>
- Scherbaum F, Delavaud E, Riggelsen C (2009) Model selection in seismic hazard analysis: an information-theoretic perspective. *Bull Seismol Soc Am* 99:3234–3247. <https://doi.org/10.1785/0120080347>
- Singh SK, Mena E, Castro R, Carmona C (1987) Empirical prediction of ground motion in Mexico City from coastal earthquakes. *Bull Seismol Soc Am* 77:1862–1867
- Singh SK, Lermo J, Domínguez T et al (1988) The Mexico earthquake of September 19, 1985—a study of amplification of seismic waves in the Valley of Mexico with respect to a hill zone site. *Earthq Spectra* 4:653–673
- Singh SK, Mena E, Castro R (1988) Prediction of peak, horizontal ground motion parameters in Mexico City from Coastal Earthquake. *Geofísica Internacional* 27:111–129
- Singh SK, Ordaz M, Anderson JG et al (1989) Analysis of near-source strong-motion recordings along the Mexican subduction zone. *Bull Seismol Soc Am* 79:1697–1717
- Singh SK, Quaaas R, Ordaz M et al (1995) Is there truly a hard rock site in the Valley of Mexico? *Geophys Res Lett* 22:481–484. <https://doi.org/10.1029/94GL03298>
- SSN (2022) Universidad Nacional Autónoma de México, Instituto de Geofísica, Servicio Sismológico Nacional (2021), Catálogo de sismos. UNAM, IGEF, SSN. <http://www2.ssn.unam.mx:8080/catalogo/>
- Taboada-Urtuzuástegui V, Martínez H, Romo MP (1999) Evaluation of dynamic soil properties in Mexico City using downhole array records. *Soils Found* 39:81–92. [https://doi.org/10.3208/sandf.39.5\\_81](https://doi.org/10.3208/sandf.39.5_81)
- Wang M, Takada T (2009) A Bayesian framework for prediction of seismic ground motion. *Bull Seismol Soc Am* 99:2348–2364. <https://doi.org/10.1785/0120080017>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.