Ranking of Ground-Motion Models (GMMs) for Use in Probabilistic Seismic Hazard Analysis for Iran Based on an Independent Data Set

Zoya Farajpour[1], Milad Kowsari[2], Shahram Pezeshk*[1], and Benedikt Halldorsson[2,3]

ABSTRACT

We apply three data-driven selection methods, log-likelihood (LLH), Euclidean distance-based ranking (EDR), and deviance information criterion (DIC), to objectively evaluate the predictive capability of 10 ground-motion models (GMMs) developed from Iranian and worldwide data sets against a new and independent Iranian strong-motion data set. The data set includes, for example, the 12 November 2017 M_w 7.3 Ezgaleh earthquake and the 25 November 2018 M_w 6.3 Sarpol-e Zahab earthquake and includes a total of 201 records from 29 recent events with moment magnitudes $4.5 \le M_w \le 7.3$ with distances up to 275 km. The results of this study show that the prior sigma of the GMMs acts as the key measure used by the LLH and EDR methods in the ranking against the data set. In some cases, this leads to the resulting model bias being ignored. In contrast, the DIC method is free from such ambiguity as it uses the posterior sigma as the basis for the ranking. Thus, the DIC method offers a clear advantage of partially removing the ergodic assumption from the GMM selection process and allows a more objective representation of the expected ground motion at a specific site when the ground-motion recordings are homogeneously distributed in terms of magnitudes and distances. The ranking results thus show that the local models that were exclusively developed from Iranian strong motions perform better than GMMs from other regions for use in probabilistic seismic hazard analysis in Iran. Among the Next Generation Attenuation-West2 models, the GMMs by Boore *et al.* (2014) and Abrahamson *et al.* (2014) perform better. The GMMs proposed by Darzi *et al.* (2019) and Farajpour *et al.* (2019) fit the recorded data well at short periods (peak ground acceleration and pseudoacceleration spectra at $T = 0.2$ s). However, at long periods, the models developed by Zafarani *et al.* (2018), Sedaghati and Pezeshk (2017), and Kale *et al.* (2015) are preferable.

KEY POINTS

- Three data-driven selection methods are used to evaluate the predictive capability of ten GMMs for Iran.
- Local models developed from Iranian strong motions perform better than GMMs from other regions.
- Deviance Information Criterion (DIC) optimizes the selection of GMMs using the Bayesian statistical framework.

Supplemental Material

INTRODUCTION

In current building codes, seismic hazard maps that are obtained by a probabilistic seismic hazard analysis (PSHA), characterize the strong ground motion for designing a structure to resist earthquakes in a given region. The purpose of PSHA is to determine the annual probability of exceedance of a ground-motion intensity measure for a seismic region of a given seismic activity and a prescribed attenuation of that measure as a function of independent variables. For earthquake engineering purposes, the most commonly used intensity measures are the recorded peak ground acceleration (PGA) and the 5% damped pseudoacceleration spectra (PSA) evaluated at discrete oscillating periods in a range of engineering interest. The peak parameters are predicted based on key independent variables such as earthquake magnitude, source–site distance, and

1. Department of Civil Engineering, The University of Memphis, Memphis, Tennessee, U.S.A.; 2. Earthquake Engineering Research Centre and Faculty of Civil and Environmental Engineering, School of Engineering and Natural Sciences, University of Iceland, Reykjavik, Iceland; 3. Geoscience Research Group, Division of Processing and Research, Icelandic Meteorological Office, Reykjavik, Iceland

*Corresponding author: spezeshk@memphis.edu

Cite this article as Farajpour, Z., M. Kowsari, S. Pezeshk, and B. Halldorsson (2020). Ranking of Ground-Motion Models (GMMs) for Use in Probabilistic Seismic Hazard Analysis for Iran Based on an Independent Data Set, *Bull. Seismol. Soc. Am.* **XX**, 1–16, doi: 10.1785/0120200052

soil amplification (site) effects, along with various other parameters. The attenuation of seismic ground motion is generally represented by empirical ground-motion models (GMMs) that quantify the salient characteristics of the propagation of seismic ground motion through the target region. Despite the modeling efforts and increasingly available data over the last couple of decades, the scatter of model residuals, quantified as the prediction uncertainty of GMMs, exerts a greater influence on the hazard analysis results than other sources of uncertainties related to the fundamental input assumptions in the assessment (Cramer, 2001; Petersen *et al.*, 2004; Bommer *et al.*, 2005; Cao *et al.*, 2005; Lombardi *et al.*, 2005; Sabetta *et al.*, 2005; Atkinson and Goda, 2011; Bradley *et al.*, 2012; Kowsari *et al.*, 2018; Kowsari, Eftekhari, *et al.*, 2019; Kowsari, Halldorsson, Hrafnkelsson, Snæbjörnsson, *et al.*, 2019).

In GMMs, the residual scatter denoted by sigma (σ) is interpreted as the aleatory variability of ground motion, whereas the uncertainty of the correct value of the median is considered to be epistemic (Strasser *et al.*, 2009; Atkinson *et al.*, 2014). The epistemic uncertainty that arises from the lack of knowledge and leads inevitably to imperfect models is generally accounted for by implementing multiple GMMs in a logic-tree framework (Kulkarni *et al.*, 1984) or by capturing the center, body, and range of ground-motion estimates in a backbone approach (Atkinson and Adams, 2013; Douglas, 2018a,b). However, the selection of appropriate GMM is still of great importance in both approaches, appropriate in the sense that the selected GMMs are representative of the ground motions in the region that the PSHA is to be carried out. This is not always guaranteed, in particular, for seismic regions for which data are scarce. In this vein, data-driven methods have been introduced and provide a systematic way to reduce subjectivity and quantitatively guide the selection process (Delavaud *et al.*, 2009, 2012; Scherbaum *et al.*, 2009). The applicability of GMMs to different regions has been widely examined using different data-driven methods (Hintersberger *et al.*, 2007; Stafford *et al.*, 2008; Shoja–Taheri *et al.*, 2010; Beauval *et al.*, 2012; Farhadi *et al.*, 2019). A likelihood-based approach (called LH) was proposed by Scherbaum, Cotton, and Smit (2004) as one of the first data-driven methods. The LH method is the exceedance probability-based approach that calculates the normalized residuals for a set of observed and estimated ground-motion data. Scherbaum *et al.* (2009) suggested an information-theoretic approach called log-likelihood (LLH) method that overcomes several shortcomings of the LH method. The LLH is less dependent on the sample size as compared with the LH method, and also it does not require any ad hoc assumptions regarding classification boundaries (Delavaud *et al.*, 2009). Another data-driven method was introduced by Kale and Akkar (2013) that uses the Euclidean distance (the absolute difference between the observed and estimated data) to account for both aleatory variability in ground motions and the trend between the observed and estimated data. Mak *et al.* (2017) represented the effects of data correlation and score variability on the evaluation of GMMs. Kowsari, Halldorsson, Hrafnkelsson, and Jonsson (2019) proposed a new data-driven method using the deviance information criterion (DIC) that optimizes the selection of GMMs using the Bayesian statistical framework.

Ideally, the data-driven selection methods should be applied to the test data that are of high quality and independent of all the evaluated models. Otherwise, the results of the evaluation of GMMs may not be reliable (Mak *et al.*, 2017). Thus, the selection method and the data set used are the two key elements that need to be carefully checked and chosen for evaluating the GMMs to be used in PSHA. In the case of Iran, it is one of the most seismically active regions of the world with the Iranian plateau located in the Alpine–Himalayan orogenic belt and wedged between Eurasia and Arabian plate stable platforms. Its level of seismicity has been manifested throughout history by repeated occurrences of destructive earthquakes. The Iranian strong-motion network (ISMN) was deployed in 1973 and currently has more than 1047 active recording stations spread all over the country, with the majority in the southwest and western part of Iran due to the higher seismicity of the Zagros belt. The high-quality strong-motion data provided by ISMN has allowed researchers to develop several GMMs for Iran (e.g., Kale *et al.*, 2015; Sedaghati and Pezeshk, 2017; Zafarani *et al.*, 2018; Darzi *et al.*, 2019; Farajpour *et al.*, 2019).

The selected database is from the Building and Housing Research Center (BHRC) seismographic stations to obtain strong ground motions of earthquakes. The recent destructive earthquakes strong-motion data such as the 12 November 2017 M_w 7.3 Ezgaleh earthquake and the 25 November 2018 M_w 6.3 Sarpol-e Zahab earthquake, which occurred in western Iran, are included. We compiled a database of horizontal ground-motion acceleration time histories for events with moment magnitudes $4.5 \le M_w \le 7.3$ at distances of 6–275 km from 2000 to 2019 for Iran.

This data set lends itself especially well for this study, namely the testing of the capability of recent GMMs to represent Iranian strong motions. In this study, therefore, we consider 10 different GMMs that have been developed based on local and/or worldwide data and have been recommended for application of PSHA in Iran. We rank these GMMs against the new and the independent data set of Iranian earthquake strong motions using three different data-driven methods, the likelihood-based (LLH), the Euclidean distance-based ranking (EDR), and a DIC. The results of this study will show which GMMs most efficiently represent the data set used and will facilitate the selection of appropriate GMMs for Iran. We believe the results of this study will constitute a significant contribution to the future seismic hazard studies in Iran by suggesting the most appropriate models to be implemented in a reliable PSHA that will result in reducing the epistemic uncertainty.

TABLE 1
List of Earthquake Events in the Data Set Used

Earthquake Date (yyyy/mm/dd)	Event_Lat (°)	Event_Lon (°)	Depth (km)	M_w	Region	Number of Records	Distance Range (km)
2000/12/06	39.53	54.81	2	6.3	Alborz and Azerbaijan	3	200–220
2012/04/18	27.85	58.10	60	5.1	Others	3	62–76
2013/04/18	38.39	45.37	12	4.9	Alborz and Azerbaijan	2	38–109
2014/08/18	32.64	47.64	12	4.5	Zagros	2	15–36
2015/07/31	29.99	57.63	10	5.4	Central Iran	11	14–80
2016/07/28	26.84	53.82	18	4.8	Zagros	3	17–40
2017/01/17	29.66	51.56	20	4.9	Zagros	2	19–27
2017/04/05	35.89	60.37	9	6.1	Kope Dagh	11	47–163
2017/05/02	35.81	60.54	10	4.8	Kope Dagh	5	37–87
2017/05/11	39.81	48.51	56	5.2	Alborz and Azerbaijan	7	75–119
2017/05/13	37.58	57.19	7	5.8	Kope Dagh	7	17–94
2017/07/23	30.09	57.62	8	5.3	Central Iran	3	6–35
2017/07/30	31.90	50.69	7	4.5	Zagros	4	7–22
2017/08/27	37.90	47.09	12	5.1	Alborz and Azerbaijan	3	10–48
2017/10/23	27.76	57.10	26	5.4	Others	3	26–50
2017/11/12	34.81	45.91	18	7.3	Zagros	40	19–275
2017/12/01	30.75	57.34	14	6.1	Central Iran	12	21–111
2017/12/20	35.77	50.90	14	4.9	Alborz and Azerbaijan	4	14–72
2018/01/06	34.47	45.79	14	5.1	Zagros	8	15–47
2018/01/11	33.78	45.76	12	5.5	Zagros	4	18–50
2018/01/11	33.70	46.01	16	4.8	Zagros	2	16–41
2018/01/11	33.86	45.85	18	5.5	Zagros	3	23–54
2018/03/19	29.712	50.78	10	4.9	Zagros	2	28–33
2018/04/19	28.35	51.64	16	5.7	Zagros	12	15–97
2018/05/02	30.80	51.66	10	5.3	Central Iran	5	21–67
2018/07/22	30.36	57.49	18	5.7	Central Iran	6	20–45
2018/08/28	38.76	48.71	24	5.1	Alborz and Azerbaijan	4	44–94
2018/11/25	34.31	45.69	16	6.3	Zagros	23	13–154
2019/01/06	34.12	45.53	10	5.7	Zagros	7	27–91

IRANIAN GROUND-MOTION DATA SET

After establishment of the ISMN in 1973, many destructive earthquakes with magnitudes larger than 6.0 have occurred in Iran such as the Sirch 1981 (M_w 6.6), Rudbar-Manjil 1990 (M_w 7.4), Ardakol 1997 (M_w 6.9), Bam 2003 (M_w 6.5), Aghgala 2004, 2005 (M_w 6.2, 6.7, respectively), Silakhor 2006 (M_w 6.2), Varzaqan 2012 (M_w 6.2), Hajdak 2017 (M_w 6.1), and Ezgaleh 2017 (M_w 7.3) earthquakes. Before 1973, there were no instrumental records but only historical evidence of damaging earthquakes in different parts of Iran. The BHRC, the agency that owns and operates the Iranian national strong-motion network, is the main source for the Iranian strong-motion records. Its investment in improving the availability of strong-motion data is an important contribution that leads to improved ground-motion prediction models that in turn are the basis for increasing the reliability of the earthquake hazard assessment in Iran. The data set used in this study is limited in terms of large magnitudes. From 2012 to 2019, 12 earthquakes with magnitude larger than M_w 6 have occurred in Iran in which six of them (from 2012 to 2014) were used in the previous studies (e.g., Darzi *et al.*, 2019; Farajpour *et al.*, 2019). To keep the data set independent, we only used strong-motion records of those large events that were available and not used in the previous studies. We primarily used the strong-motion data from the Ezgaleh and Sarpol-e Zahab earthquakes in western Iran along with other moderate-to-large earthquakes that recently occurred in the northwestern part of the Zagros fold-and-thrust belt (ZFTB). These earthquakes, along with the other shallow crustal earthquakes recorded between 2000 and 2019, have provided a high-quality strong-motion data set including 201 records with moment magnitudes $4.5 \le M_w \le 7.3$ and distances up to 275 km making this data set appropriate for testing the applicability of different GMMs for use in PSHA in Iran and are provided in Table 1. The flat file of the data set used is provided in the supplemental material to this article. The geographic distribution of the used recording stations and earthquake locations are shown in Figure 1.

The ZFTB mountain ranges extend for ~1500 km from the western part of Iran to northern Iraq. The plate-motion model estimates a 30 mm/yr collision rate for ZFTB (DeMets *et al.*, 1994). This area has frequently experienced devastating

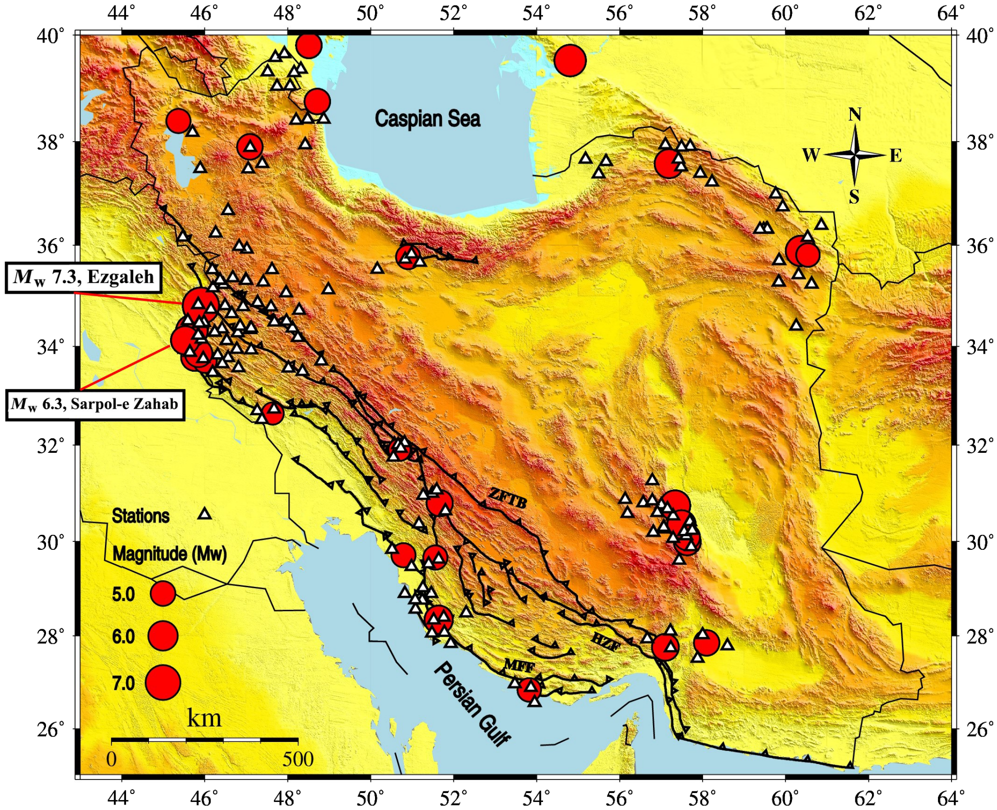

Figure 1. Geographic distribution of Iranian strong-motion recording stations (triangles) and earthquakes epicenters (circles) used in this study. The Zagros fold-and-thrust belt (ZFTB), high Zagros fault (HZF), and Mountain Front fault (MFF) are also shown. The colors indicate height over sea level with red being the mountainous areas. The color version of this figure is available only in the electronic edition.

earthquakes, mostly shallower than 15 km depth. Moreover, the high Zagros fault (HZF) and the Mountain Front fault (MFF) are two major active thrust fault systems in the northwest of Zagros. One of these major faulting systems is very likely to be responsible for the Sarpol-e Zahab event (Chen *et al.*, 2018). According to Engdahl *et al.* (2006), thrust-faulting events occur at depths between 10 and 20 km in this region. However, the hypocenter of the Sarpol-e Zahab earthquake is located beneath the HZF at ~15 km depth, and it can rule out the HZF to be the source fault. Thus, it was a blind oblique-thrust faulting event on the deep section of the MFF (Chen *et al.*, 2018).

In this study, we have compiled and processed the strong-motion data following the approach used by Farajpour *et al.* (2018). The data set is processed by the first baseline offset correcting the time histories by following the approach recommended by Boore *et al.* (2002). Then bandpass filtering techniques were applied on the strong-motion data with corner frequencies often chosen based on the shape of Fourier amplitude spectra and the signal-to-noise ratio (Boore and Bommer, 2005). We removed records with signal-to-noise ratio lower than 3 to avoid the confusion of choosing acausal (phase-less) filter frequencies (Farajpour *et al.*, 2018). The data used are the PGA and 5% damped PSA at different periods. Two horizontal components of the ground motions are combined into a single measure using the geometric mean, which is one of the accepted and most used measures in GMMs.

The magnitude scale of the strong-motion data set is the moment magnitude that is reported and estimated by BHRC. Moreover, the data set includes different source-to-site distances such as epicentral distance, hypocentral distance, rupture distance ($R_{\rm rup}$), and the Joyner–Boore distances ($R_{\rm JB}$). For site effects, the average shear-wave velocity estimates in the upper 30 m (V_{S30}) were used. Some of the selected stations have V_{S30} values, which were measured by a shallow seismic refraction technique. In this study, for those stations without measured V_{S30}, we used the topographic slope as a proxy to V_{S30} (Wald and Allen, 2007). A similar approach was used by the Next Generation Attenuation-West2 (NGA-West2) and the NGA-East to estimate V_{S30} for sites where there are no measured data. Moreover, many of the Iranian sites were constructed on rock, so the distribution is biased toward the higher-velocity V_{S30} data (Karimzadeh *et al.*, 2019). Thus, the data set used includes 111 records with measured V_{S30} and 90 records with estimated values. Sites are categorized based on the National Earthquake Hazards Reduction Program 2000 (Federal Emergency Management Agency-368, 2000) site-class classification. Figure 2 shows the magnitude versus distance distributions and V_{S30} distributions of the selected records.

It is important to mention that the data set used in this study was not used by the GMM developers and is an independent data set. Figure 3 shows a comparison between the data set used in this study with the strong-motion data set prepared by Farajpour *et al.* (2018), shown in gray for PGA and PSA at $T = 0.2$, 1.0, and 2.0 s. The diamonds and circles show the ground-motion intensity measure versus magnitude and distance, respectively. Figure 3 indicates that the independent data set used is generally consistent with the characteristics of the Iranian strong motions.

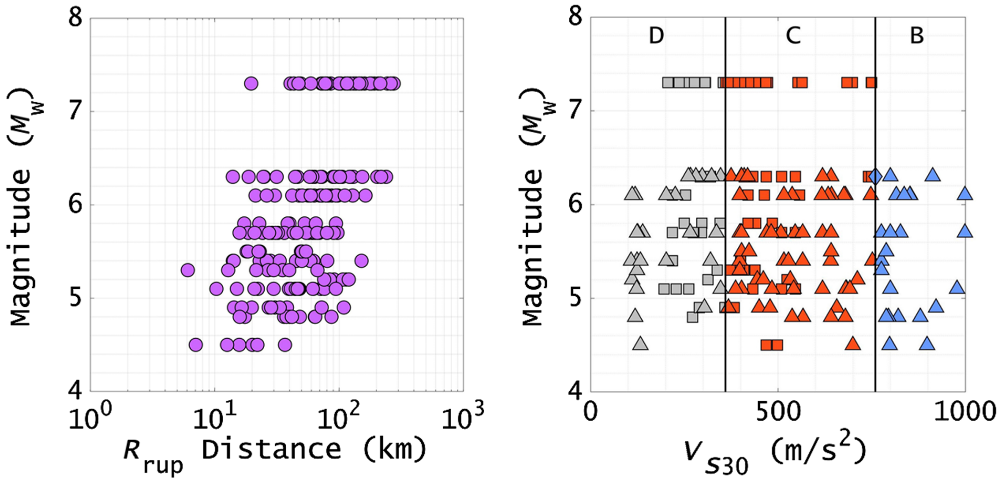

Figure 2. The characteristics of the earthquake strong-motion data set used in this study in terms of magnitude–distance distribution (left) and the distribution of V_{S30} of the recording sites with magnitude. The site categories based on the National Earthquake Hazards Reduction Program 2000 (Federal Emergency Management Agency-368, 2000) and their relation to V_{S30} are also shown (B, 760 ≤ V_{S30} < 1500 m/s; C, 360 ≤ V_{S30} < 760 m/s; D, 180 ≤ V_{S30} < 360 m/s). The sites with measured values (triangles) and the estimated sites (squares) are also shown. The color version of this figure is available only in the electronic edition.

SELECTED GMMS

Ten empirical GMMs are considered in this study, all of which satisfy the minimum requirements proposed by Cotton *et al.* (2006) and Bommer *et al.* (2010). Five GMMs are recent GMMs that were developed exclusively using the Iranian strong-motion data, and the rest are the GMMs developed from strong-motion data worldwide. GMMs considered fall into two categories: (1) the local models including Farajpour *et al.* (2019, hereafter, FPZ19), Darzi *et al.* (2019, hereafter, Dea19), Zafarani *et al.* (2018, hereafter, Zea18), Sedaghati and Pezeshk (2017, hereafter, SP17), Kale *et al.* (2015, hereafter, Kea15); (2) the NGA models (NGA-West2) including Abrahamson *et al.* (2014, hereafter, ASK14), Boore *et al.* (2014, hereafter, BSSA14), Chiou and Youngs (2014, hereafter, CY14), Campbell and Bozorgnia (2014, hereafter, CB14), and Idriss (2014, hereafter, I14). Table 2 lists the GMMs considered in this study and their range of applicability, combination method for the two horizontal components, and region of origin.

The FPZ19 is an empirical model for the prediction of PGA and 5% damped PSA up to 4.0 s. The model is based on a data set with 1356 strong-motion records from 208 events with a

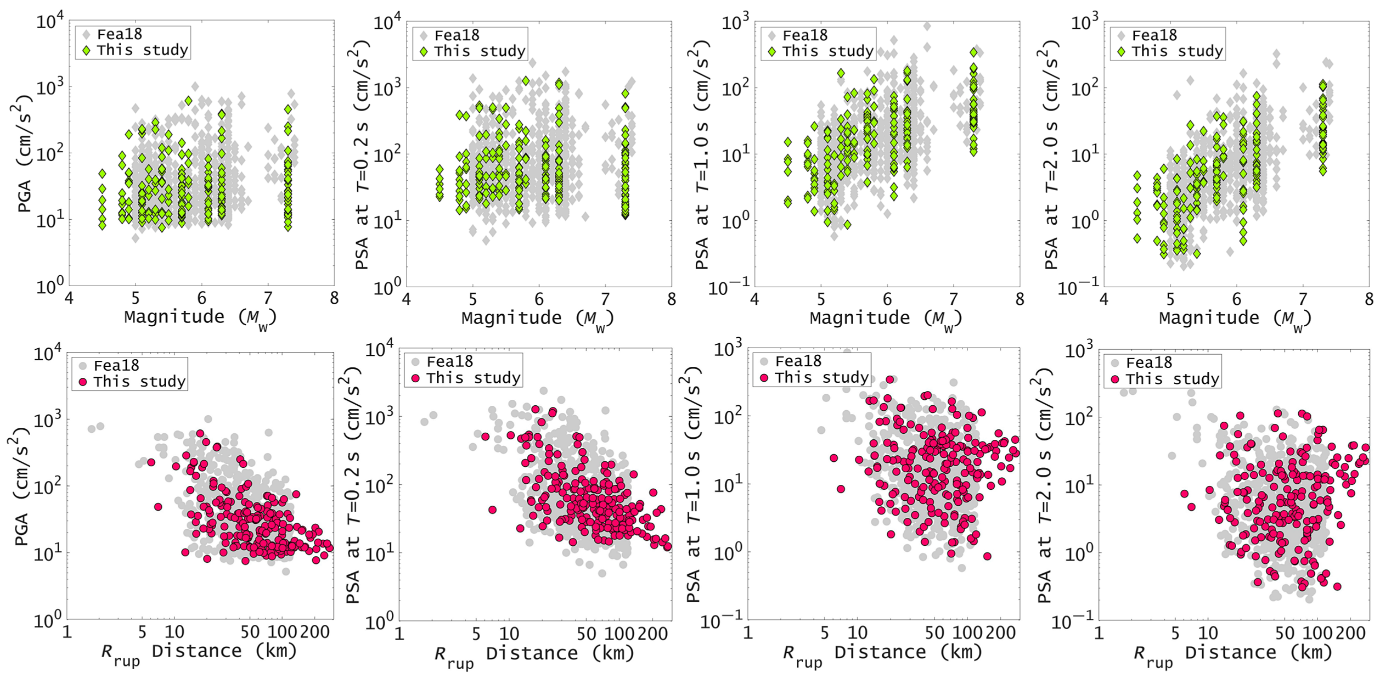

Figure 3. The comparison of the independent strong-motion data set used in this study (shown in diamonds and circles) with the data set prepared by Farajpour *et al.* (2018, hereafter, Fea18) (shown in gray). The peak ground acceleration (PGA) and pseudoacceleration spectra (PSA) at $T = 0.2$, 1.0, and 2.0 s (columns left to right) are shown versus magnitude (top row) and distance (bottom row). The color version of this figure is available only in the electronic edition.

TABLE 2
Description of the Earthquake Ground-Motion Models (GMMs) Used in This Study

GMM	*M* Range	*R* Type	*R* Range (km)	Horizontal Component	Main Region(s)
FPZ19 (Farajpour *et al.*, 2019)	4.8–7.5	R_{rup}	0–400	GM	Iran
Dea19 (Darzi *et al.*, 2019)	4.5–7.4	R_{JB}, R_{rup}	0–200	GM	Iran
Zea18 (Zafarani *et al.*, 2018)	4.0–7.3	R_{JB}, R_{epi}	0–200	GM	Iran
SP17 (Sedaghati and Pezeshk, 2017)	4.7–7.4	R_{JB}	0–250	GM	Iran
Kea15 (Kale *et al.*, 2015)	4.0–8.0	R_{JB}	0–200	GM	Turkey and Iran
ASK14 (Abrahamson *et al.*, 2014)	3.0–8.5	R_{JB}, R_{rup}	0–300	RotD50	Worldwide
BSSA14 (Boore *et al.*, 2014)	3.0–8.5	R_{JB}	0–400	RotD50	Worldwide
CY14 (Chiou and Youngs, 2014)	3.5–8.5	R_{JB}, R_{rup}	0–300	RotD50	Worldwide
CB14 (Campbell and Bozorgnia, 2014)	3.3–8.5	R_{JB}, R_{rup}	0–300	RotD50	Worldwide
I14 (Idriss, 2014)	4.5–7.9	R_{rup}	0–175	RotD50	Worldwide

GM is the geometric mean and RotD50 is the 50th percentile of the rotated orientation-independent.

magnitude range of 4.8–7.5, and the rupture distances (R_{rup}) up to 400 km. This GMM considered regional differences for five major tectonic regions and included the nonlinear site effect using V_{S30}. The Dea19 is another newly developed empirical model for the prediction of peak ground velocity (PGV), PGA, and 5% damped PSA up to 10 s. The model is based on 1350 records from 370 earthquake events with a magnitude range of 4.5–7.4 and source-to-site distances up to 200 km. They used an adaptive wavelet denoising approach for waveform processing. In addition, they considered regional differences for different tectonic regimes area, and significant regional differences were reported for a few magnitude–distance intervals. The Zea18 is derived from 1551 Iranian acceleration to predict PGA and 5% damped PSA up to 4.0 s from 200 shallow earthquakes distances up to 200 km. The SP17 model estimates horizontal and vertical strong ground motion intensity measures for Iran's shallow crustal earthquakes. Their data set includes 688 records from 152 earthquakes with moment magnitudes ranging from 4.7 to 7.4 and Joyner–Boore distances up to 250 km. The local site conditions were also considered in their model by accounting the averaged shear-wave velocities in the upper 30 m. Kea15 is derived from a subset of 670 Turkish and 528 Iranian accelerograms for shallow earthquake event with depth less than 35 km to predict PGA, PGV, and 5% damped PSA.

The NGA-West1 empirical GMMs (Power *et al.*, 2008) had been applied in previous seismic hazard studies in Iran, where it had been assumed that they adequately describe the characteristics of Iranian strong motions (e.g., Khoshnevis *et al.*, 2017). NGA-West2 models were developed by five research teams from the Pacific Earthquake Engineering Research center for the prediction of seismic wave attenuation from shallow crustal earthquakes in interplate regions. Some researchers believe that these models are more reliable than local GMMs for estimating the ground motion in the region where PSHA is going to be carried out. However, using fully ergodic GMMs is effectively equivalent to applying the aleatory variability from other regions to the region under study, potentially introducing unrealistic uncertainties into PSHA results (Kowsari, Halldorsson, Hrafnkelsson, Snæbjörnsson, *et al.*, 2019). The NGA-West2 models were derived from a data set of worldwide strong-motion recordings, in which Iranian strong-motion data were rarely used. The developers have used different data selection criteria, parameters, and functional forms for their models (e.g., Bozorgnia *et al.*, 2014). Particularly, the applied functional forms allow for more flexibility in the modeling of the ground-motion scaling with magnitude from moderate-to-large earthquakes, and the near-fault magnitude-dependent saturation of amplitudes with distance. Nevertheless, their applicability for use in PSHA in Iran is a matter of question that this study tries to address.

RANKING GMMS USING DATA-DRIVEN METHODS

The LLH method uses the Kullback–Leibler metric (i.e., the distance between the expected value of the correct model and the predicted value of the approximate model). Practically, LLH measures the distance between two continuous probability density functions, such as $f(x)$ and $g(x)$. The function $f(x)$ is the log-normal distribution function of the observed data point (x_i), and $g(x_i)$ are estimated values that are distributed log-normally with the median and standard deviation of the considered GMM. The LLH score for N pairs of observation and prediction is calculated by the following equation:

$$\mathrm{LLH}(g,x) = \frac{-1}{N}\sum_{i=1}^{N}\log_2(g(x_i)). \quad (1)$$

EDR is another ranking method used in this study. The methodology modifies the concept of the Euclidean distance and separately considers the standard deviation of the GMMs and the bias between the observed data and median estimations. The DE is similar to the residual analysis concept, which is given as

$$\mathrm{DE} = \sqrt{\sum_{i=1}^{N} (p_i - q_i)^2}, \tag{2}$$

in which N is the total number of observations x_i, the p_i and q_i parameters are the observed and the predicted ground-motion pairs, respectively. The modified Euclidean distance (MDE) is calculated for a preselected standard deviation range. It modifies the median estimations with the straight line fitted to the observed and estimated ground-motion data, and its formula is given by

$$\mathrm{MDE} = \sum_{j=1}^{n} |d_j| \Pr(|\mathrm{D}| < |d_j|), \tag{3}$$

in which D is the difference between natural logarithms of the observed and the predicted ground motions and d_j represents the discrete values of D. This method considers only positive values because of the analogy made between DE and D (see Kale and Akkar, 2013, for more details). In the EDR method, the parameter k is introduced as a ratio of the original and the corrected Euclidian distance that consider the bias for the GMMs:

$$k = \frac{\mathrm{DE}_{\mathrm{original}}}{\mathrm{DE}_{\mathrm{corrected}}}. \tag{4}$$

Finally, the MDE values are combined with k, and normalized by the total numbers of data to present the final EDR ranking index:

$$\mathrm{EDR} = \sqrt{k \times \frac{1}{N} \times \sum_{i=1}^{N} \mathrm{MDE}_{\mathrm{i}}^2}. \tag{5}$$

The smallest value of EDR would show the best goodness of fit of the GMM model to the data set.

Kowsari, Halldorsson, Hrafnkelsson, and Jonsson (2019) proposed the data-driven method using the DIC to select the most suitable GMM for application in PSHA. The aleatory variability that is represented by the standard deviation of the GMM is an influential parameter in data-driven methods. To have a reliable PSHA, a nonergodic or partially nonergodic GMMs should be applied. In this way, two cases were assumed for the modeling in their study: (1) The standard deviation of the model is assumed to be known (i.e., the previously determined sigma that hereafter is denoted as prior sigma), and (2) the standard deviation of the model is unknown. They showed that the former case (i.e., the prior sigma) and the LLH method similarly ranks the GMMs because they are identical because they both use the Kullback–Leibler divergence estimated from the statistical expectation of LLH of observations. The Kullback–Leibler divergence is a measure of the difference between two probability distributions (i.e., the distribution of ground-motion observations and predictions). The only difference that has made the DIC beneficial is how the DIC method can be connected to the Bayesian statistics and the Markov chain Monte Carlo (MCMC) algorithm, which is modeled in case 2 where a posterior sigma is estimated by the Bayesian approach. That way, the posterior sigma represents the misfit between the predicted ground motions and the Local or regional ground-motion observations of the region where PSHA is going to be carried out.

The Bayesian statistical framework is ideal for incorporating in a quantifiable way our prior knowledge of model parameters and their uncertainties, which presumably we can update when new information becomes available and repeat the analysis. The output is known as a posterior distribution, which might be used as the basis for inferential decisions. Therefore, such a useful feature can be used to estimate the standard deviation of GMMs by combining a prior distribution with the likelihood of the available data. The posterior probability distribution of the unknown sigma is then conditioned on the observed ground motions obtained from the region under study. Here, we assume that the logarithm of the ground-motion parameter follows a normal distribution (Kowsari, Halldorsson, Hrafnkelsson, and Jonsson, 2019):

$$p(y|\sigma^2) = \prod_{i=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y_i - \mu(\beta)_i)^2}{2\sigma^2}\right), \tag{6}$$

in which N is the number of observations, y is the natural logarithm of the observed ground-motion parameter, $\mu(\beta)$ is the mean value predicted by the GMM, and σ is the standard deviation of the GMM. On the other hand, for the prior distribution, we assume that sigma is unknown and distributed following a scaled inverse chi-squared distribution (Kowsari, Halldorsson, Hrafnkelsson, and Jonsson, 2019):

$$p(\sigma^2) \propto (\sigma^2)^{-\frac{\nu}{2}-1} \exp\left(-\frac{\nu}{2} \times \frac{s^2}{\sigma^2}\right), \tag{7}$$

in which v is the number of chi-squared degrees of freedom, and s^2 is the scaling parameter. Therefore, the posterior distribution is given by

$$p(\sigma^2|y) \propto p(\sigma^2) p(y|\beta, \sigma^2). \tag{8}$$

By taking the logarithm of both sides:

$$\begin{aligned} \log p(\sigma^2|y) &= \log p(\sigma^2) + \log p(y|\beta, \sigma^2) \\ &= \left(-\frac{N+\nu}{2} + 1\right)\log(\sigma^2) \\ &- \frac{(N+\nu)}{2\sigma^2}\left[\frac{1}{N+\nu} + \left\{\nu s^2 + \sum_{i=1}^{N}(y_i - \mu(\beta_i))^2\right\}\right] + c \\ &\rightarrow p(\sigma^2|y) = \left(\sigma^2|N+\nu, \frac{1}{N+\nu}\right)\left\{\nu s^2 + \sum_{i=1}^{N}(y_i - \mu(\beta)_i)^2\right\}. \end{aligned} \tag{9}$$

The posterior distribution of the variance is again a scaled inverse chi-squared distribution with the degrees of

TABLE 3
Prior and Posterior Standard Deviations (in Natural Logarithm) of the Candidate GMMs at Different Periods

	Sigma (PGA)		Sigma ($T = 0.2$ s)		Sigma ($T = 1.0$ s)		Sigma ($T = 2.0$ s)	
GMMs	**Prior**	**Posterior**	**Prior**	**Posterior**	**Prior**	**Posterior**	**Prior**	**Posterior**
FPZ19	0.753	0.693	0.846	0.698	0.867	0.744	0.878	0.852
Zea18	0.686	0.637	0.751	0.725	0.787	0.730	0.781	0.760
Dea19	0.608	0.631	0.638	0.731	0.640	0.771	0.688	0.820
SP17	0.539	0.661	0.646	0.723	0.781	0.711	0.808	0.741
Kea15	0.637	0.745	0.546	0.909	0.645	0.698	0.775	0.761
ASK14	0.660	0.686	0.697	0.817	0.751	0.741	0.769	0.824
BSSA14	0.605	0.734	0.619	0.825	0.692	0.712	0.700	0.770
CY14	0.618	0.736	0.683	0.810	0.686	0.742	0.662	0.799
CB14	0.646	0.981	0.716	1.171	0.738	0.721	0.724	0.806
I14	0.763	0.749	0.808	0.839	0.833	0.839	0.843	0.872

PGA, peak ground acceleration.

freedom and the scaling parameter. As a result, the DIC of the normal model is given by (for more details see Kowsari, Halldorsson, Hrafnkelsson, and Jonsson, 2019):

$$\mathrm{DIC} = \frac{2}{L}\sum_{l=1}^{L}\left(N\log(2\pi) + N\log(\hat{\sigma}^2) + \sigma^{-2}\sum_{i=1}^{N}(y_i - \mu(\beta)_i^2)\right.$$
$$\left. - \left(N\log(2\pi) + N\log(\bar{\sigma}^2)\right) + \bar{\sigma}^2\sum_{i=1}^{N}(y_i - \mu(\beta)_i)^2\right), \quad (10)$$

in which L is the number of samples, $\hat{\sigma}^2$ are drawn samples from the posterior distribution (equation 9) using an MCMC method, $\bar{\sigma}^2$ and is the posterior mean value of sigma-squared. Therefore, the DIC is one of the model selection methods that are particularly useful where MCMC simulations to obtain the posterior distributions of the models.

RESULTS AND DISCUSSION

In this study, the performance of five recently developed local (Iranian) GMMs are evaluated using each of the three data-driven methods. Moreover, the applicability of the NGA models for use in PSHA in Iran that has been a matter of ongoing discussion among researchers is assessed. The results are presented for PGA and PSA at $T = 0.2$, 1.0, and 2.0 s. The DIC results are presented for two cases: (1) where the original sigma (the prior sigma) is used, and (2) where the sigma is assumed to be unknown and follows a scaled inverse chi-squared distribution. The prior and posterior sigma (estimated by equation 9) of the candidate GMMs are provided in Table 3. From Table 3, the posterior sigma represents the misfit between the predicted and observed ground motions of the region under study. Table 4 and Figure 4 provide scores of the EDR, LLH, and DIC with the prior sigma (named as DIC1) and DIC with the posterior sigma (named as DIC2) for the candidate GMMs at different periods.

We note that models with smaller scores in Table 4 and Figure 4 perform better compared with models with higher scores. Each ranking method is represented by different numerical values and units; however, only the relative difference within the same ranking method matters. Moreover, the results show that the DIC with the prior sigma is the same as the LLH method but with different score values, which was also shown in Kowsari, Halldorsson, Hrafnkelsson, and Jonsson (2019). The results show that all the ranking methods rank Dea19 as the best GMM at PGA. The LLH, DIC1, and DIC2 select Zea18 as the second model, whereas the EDR favors ASK14 due to its smaller sigma compared with Zea18. This can also be seen for FPZ19, where the EDR ranks it as the ninth model, whereas the other methods rank it as the fifth GMM. For PSA at $T = 0.2$ s, the EDR ranks Dea19 as the best model, whereas the other methods choose FPZ19. The Zea18 is in the second and third place by LLH and DIC2, respectively, whereas the EDR ranks it as the seventh model due to its larger sigma compared with the other GMMs. At longer periods (i.e., $T = 1.0$ and 2.0 s), Kea15 and SP17 are selected as the best GMM by the LLH, DIC1, and DIC2, whereas BSSA14 is the favored GMM by EDR. Overall, at short periods (i.e., PGA and PSA at $T = 0.2$ s) the local GMMs such as Dea19, FPZ19, Zea18, and SP17 have better performance compared with the other models, but at long periods (i.e., $T = 1.0$ and 2.0 s), Dea19 and FPZ19 lose their predictive ability compared with Zea18 and SP17. In general, the results of EDR are different from that of the LLH and DIC, and it is biased toward GMMs with smaller sigma. The behavior of NGA-West2 models is fairly similar, whereas ASK14 and BSSA14 show a better performance.

To facilitate a visual comparison of the results of the different data-driven methods, the residual trend versus magnitude and distance for the local and NGA-West2 GMMs at PGA are shown in Figures 5 and 6, respectively. For the sake of space, we show the distribution of residuals for PSA at $T = 0.2$, 1.0, and

TABLE 4

Scores of the Euclidean Distance-Based Ranking (EDR), Log Likelihood (LLH), and DIC1 with the Prior Sigma, and DIC2 with the Posterior Sigma in Selected Periods for 10 Candidate GMMs

Periods	Methods	FPZ19	Zea18	Dea19	SP17	Kea15	ASK14	BSSA14	CY14	CB14	I14
PGA	EDR	1.08	0.992	0.817	0.912	1.031	0.857	0.865	0.86	1.236	0.949
	Rank	9	7	1	5	8	2	4	3	10	6
	LLH	1.521	1.397	1.378	1.508	1.652	1.496	1.654	1.644	2.344	1.624
	Rank	5	2	1	4	8	3	9	7	10	6
	DIC1	423.9	389.4	383.9	420.1	460.2	417	460.9	458.1	653.1	452.6
	Rank	5	2	1	4	8	3	9	7	10	6
	DIC2	422.9	389	385.4	403.9	452	418.9	446.5	447.1	562.9	454.4
	Rank	5	2	1	3	8	4	6	7	10	9
$T = 0.2$ s	EDR	1.138	1.135	0.937	1.072	1.193	1.002	0.982	0.965	1.532	1.081
	Rank	8	7	1	5	9	4	3	2	10	6
	LLH	1.571	1.579	1.617	1.591	2.436	1.77	1.903	1.782	2.757	1.789
	Rank	1	2	4	3	9	5	8	6	10	7
	DIC1	437.7	440.1	450.5	443.3	678.7	493.2	530.2	496.6	768.2	498.4
	Rank	1	2	4	3	9	5	8	6	10	7
	DIC2	425.8	441.5	444.8	440.3	532.4	489.4	493.1	486.1	634.2	499.9
	Rank	1	3	4	2	9	6	7	5	10	8
$T = 1.0$ s	EDR	1.258	1.086	1.006	1.066	0.931	0.965	0.901	0.93	0.96	1.172
	Rank	10	8	6	7	3	5	1	2	4	9
	LLH	1.647	1.595	1.72	1.561	1.531	1.61	1.55	1.618	1.57	1.787
	Rank	8	5	9	3	1	6	2	7	4	10
	DIC1	458.9	444.3	479.4	434.9	426.6	448.5	432	450.9	437.5	498.1
	Rank	8	5	9	3	1	6	2	7	4	10
	DIC2	452	443.8	466.3	433.2	426.2	450.2	433.7	450.5	439.1	500
	Rank	8	5	9	2	1	6	3	7	4	10
$T = 2.0$ s	EDR	1.37	1.017	1.015	1.009	1.012	1.09	0.941	0.974	1.02	1.236
	Rank	10	6	5	3	4	8	1	2	7	9
	LLH	1.811	1.646	1.801	1.62	1.648	1.768	1.675	1.772	1.746	1.844
	Rank	9	2	8	1	3	6	4	7	5	10
	DIC1	504.7	458.8	501.9	451.3	459.2	492.8	466.8	493.6	486.5	513.9
	Rank	9	2	8	1	3	6	4	7	5	10
	DIC2	506.3	460.3	490.8	450.2	461	493	465.3	480.3	484	515.4
	Rank	9	2	7	1	3	8	4	5	6	10

DIC, deviance information criterion.

2.0 s versus distance and magnitude in the supplemental material. In these figures, the mean and the standard deviation of error bars are calculated using magnitude bins of 0.1. The solid line is the least-squares linear regression line, and the dashed lines are the 95% confidence limits of the line. However, we must emphasize that the data set used in this study is limited and only has two earthquakes with $M_w > 6.1$. Thus, the variability quantified for the large earthquakes (i.e., M_w 6.3 and 7.3) is a within-event variability.

In some cases, the performances of the EDR and LLH methods have exhibited shortcomings, which can be seen from the results presented in Table 4 and Figure 4. The LLH prefers the predictive model with a larger sigma when the observed data congregate away from the median estimations (Kale and Akkar, 2013). On the other hand, the EDR method favors a smaller sigma when two predictions give the same mean, regardless of what the true uncertainty is (Mak *et al.*, 2014). The results presented for PGA show that the EDR favors CY14 over Zea18 due to its smaller sigma (provided in Table 3 as prior sigma). From Figures 5 and 6, it is clear that the CY14 model is more biased than the Zea18. The LLH and DIC2 behave similarly where the data are close to the predictions. However, where observations are away from predictions, the LLH favors the model with a larger sigma. For example, the LLH method selects I14 over BSSA14 because of its larger sigma, although it is clear that the BSSA14 is less biased than I14 versus magnitude and distance.

At PGA, the model residual behavior versus magnitude and distance for the category 1 GMMs show that Dea19 and Zea18 are the unbiased models, meaning that the overall mean value of the residuals is insignificantly different from zero and does not show significant trends with distance or magnitude. The other local GMMs (i.e., FPZ19, SP17, and Kea15) have substantial negative trends with distance. In other words, they are underpredicting in the near-fault and overpredicting in the far field. Moreover, the local GMMs underpredict the

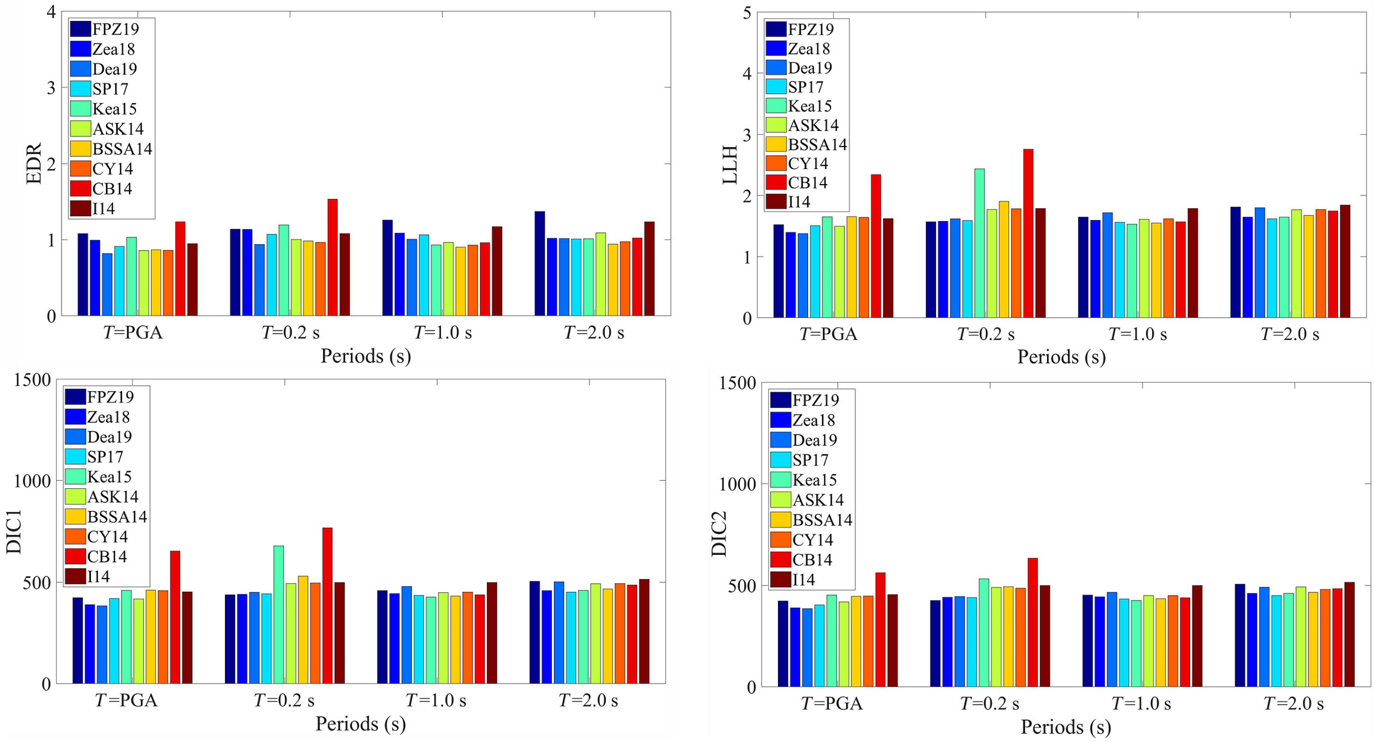

Figure 4. Euclidean distance-based ranking (EDR), log-likelihood (LLH), and DIC1 scores with the prior sigma, and DIC2 score with the posterior sigma at selected periods for the candidate ground-motion models (GMMs). The smaller scores imply better representation of the observed ground motions by the predictive model. DIC, deviance information criterion. The color version of this figure is available only in the electronic edition.

PGA of small earthquakes and overpredict for the large earthquakes, with the exception of Zea18 and Dea19. For the NGA-West2 models, the residual versus magnitude shows that the only unbiased models are BSSA14 and ASK14, although the individual event residuals for BSSA14 show large suspicious variations between M_w 5 and 6. All others have significant trends with distance and magnitude. CB14, in particular, is biased, as evident by the offset of the mean residual value from zero.

Therefore, the results indicate that the LLH and EDR methods consider the prior sigma as a key measure to rank the GMMs, and in some cases, it even leads to ignoring the model bias (represented by the deviation of the median). We believe that the prior sigma is not an appropriate measure for this purpose; therefore, we suggest using DIC2 method that ranks GMMs based on both residual and the posterior sigma obtained from the observed ground motions.

Furthermore, we plot the attenuation of categories 1 and 2 GMMs versus distance and compare them with the observed data in Figures 7 and 8 for PGA and PSA at $T = 1.0$ s, respectively. For the sake of consistency, the different distance measures were converted to the rupture distance (R_{rup}) using the relationships proposed by Scherbaum, Schmedes, and Cotton (2004). The corresponding figures for PSA at $T = 0.2$ and 2.0 s are shown in the supplemental material. In these figures, the GMMs are evaluated at M_w 5 (events with $4.5 \le M_w < 5.5$ are shown), M_w 6 (events with $5.5 \le M_w < 6.5$ are shown), and 7.3 (2017 M_w 7.3 Ezgeleh earthquake for $V_{S30} = 360$ m/s). The gray circles show the observed strong-motion data. We also plot the median GMMs that are ranked as the best model using DIC2 ± its standard deviation at each period. The results show that the 68% confidence interval (median $\pm\,\sigma$) of the first ranked GMMs fairly cover the recorded strong motions for different magnitudes at different periods. At PGA, CB14 is underpredicting the Iranian strong motions for all magnitudes, BSSA14 (slightly), and CY14 is underpredicting the small earthquakes and overpredicting the large ones as they are already shown in Figure 6. It also shows how the NGA-West2 (apart from BSSA14) overpredicts the small earthquakes at near-fault distances at PSA at $T = 1.0$ s that is compatible with Figure S4. Overall, category 1 GMMs fit the Iranian strong motions better and would be promising candidates over NGA-West2 models for use in PSHA in Iran. However, the NGA-West2 models can still be applied within Iran, which was already shown in the previous studies (Shoja-Taheri *et al.*, 2010; Mousavi *et al.*, 2012).

CONCLUSIONS

We have had a few large earthquake occurrences with magnitude of 6 and larger in Iran in recent years, which have caused significant human and financial losses. Performing a reliable seismic hazard analysis is crucial for such an earthquake-prone

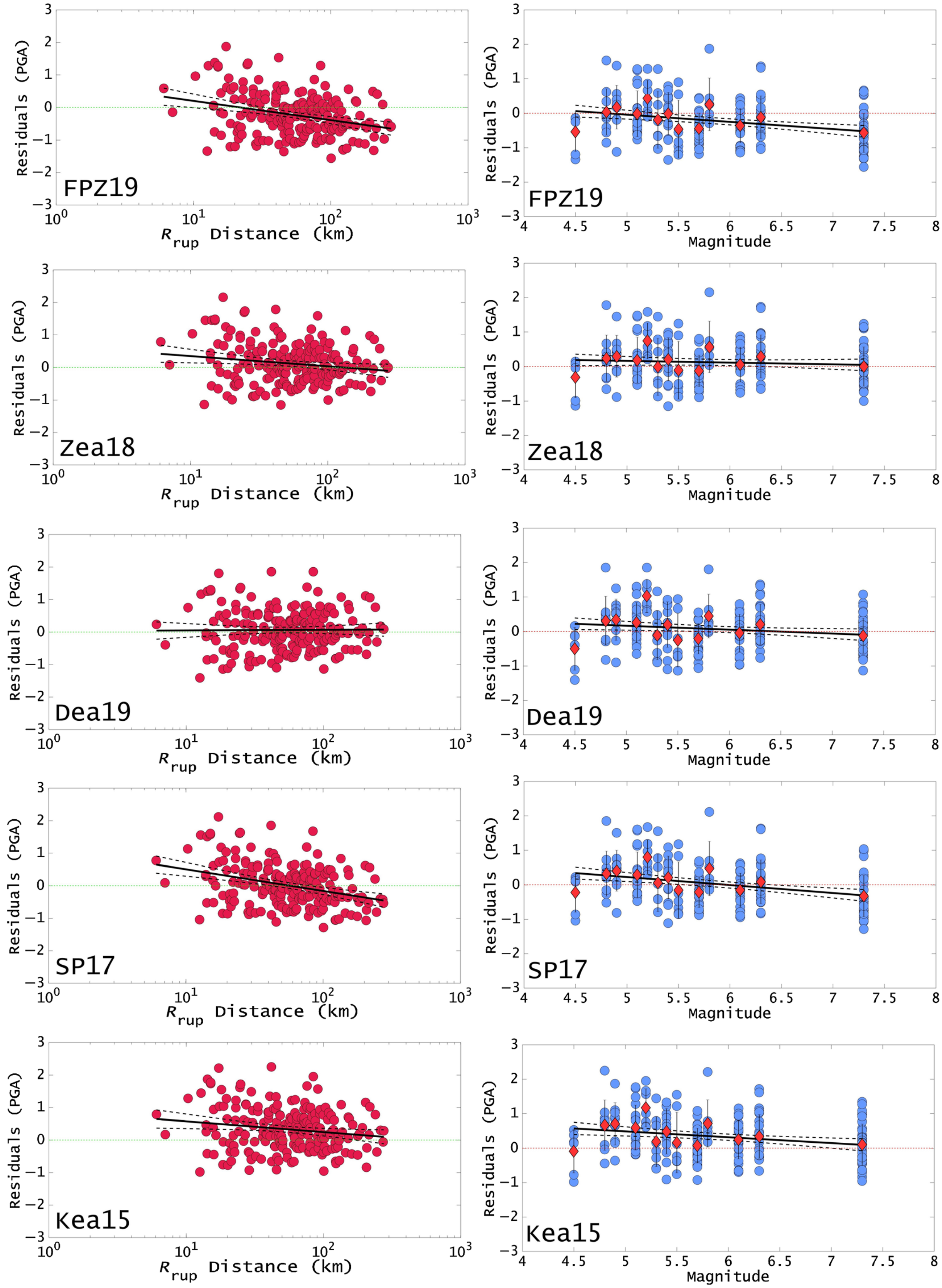

Figure 5. Distribution of residuals for PGA versus distance and magnitude for the local GMMs. Red diamonds and the error bars show the mean and standard deviation of the log residual, respectively. Solid and dashed lines are the least-squares linear regression and the 95% confidence limits, respectively. The color version of this figure is available only in the electronic edition.

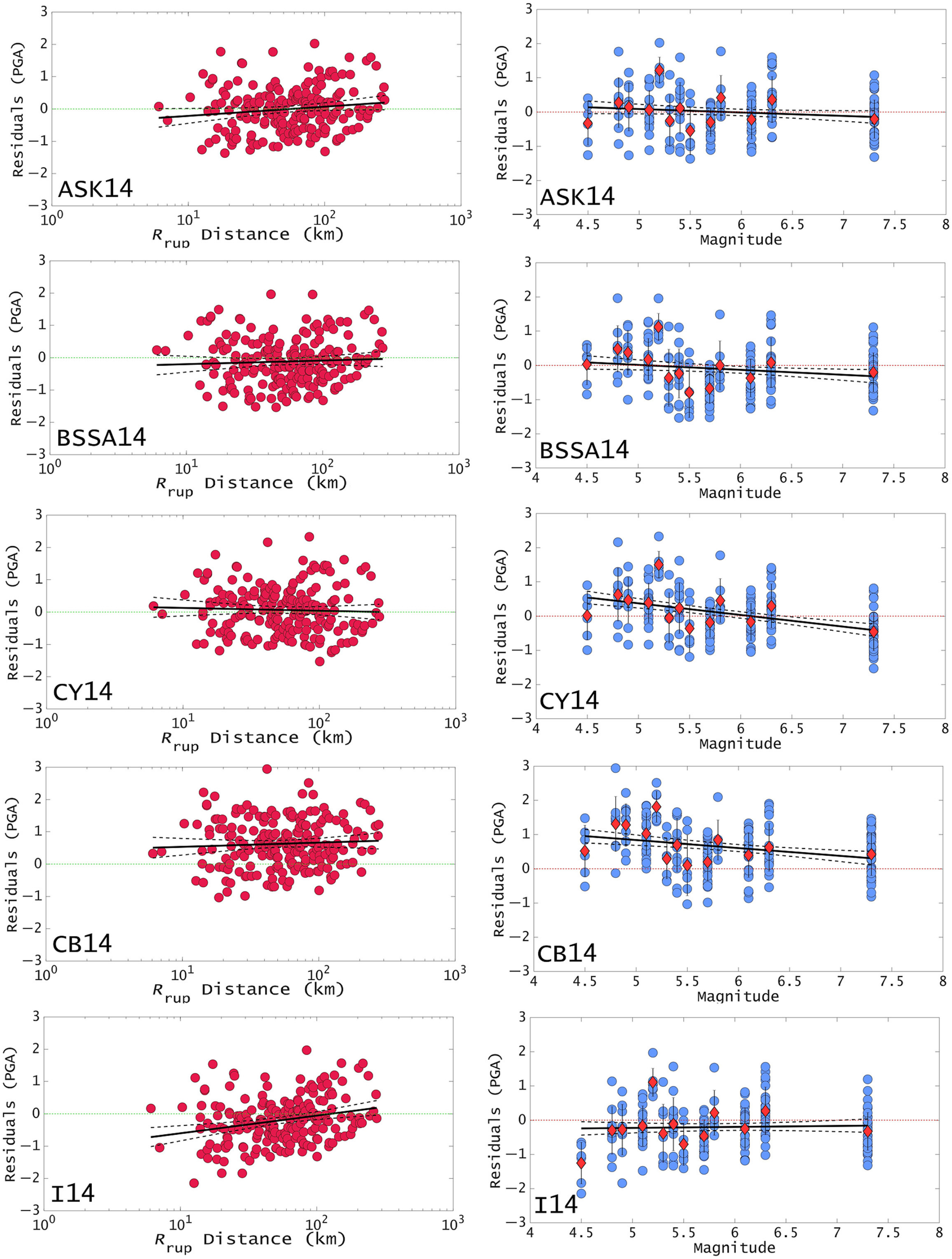

Figure 6. Distribution of residuals for PGA versus distance and magnitude for the Next Generation Attenuation-West2 (NGA-West2) GMMs. Red diamonds and error bars show the mean and standard deviation of the log residual, respectively. Solid and dashed lines are the least-squares linear regression and the 95% confidence limits, respectively. The color version of this figure is available only in the electronic edition.

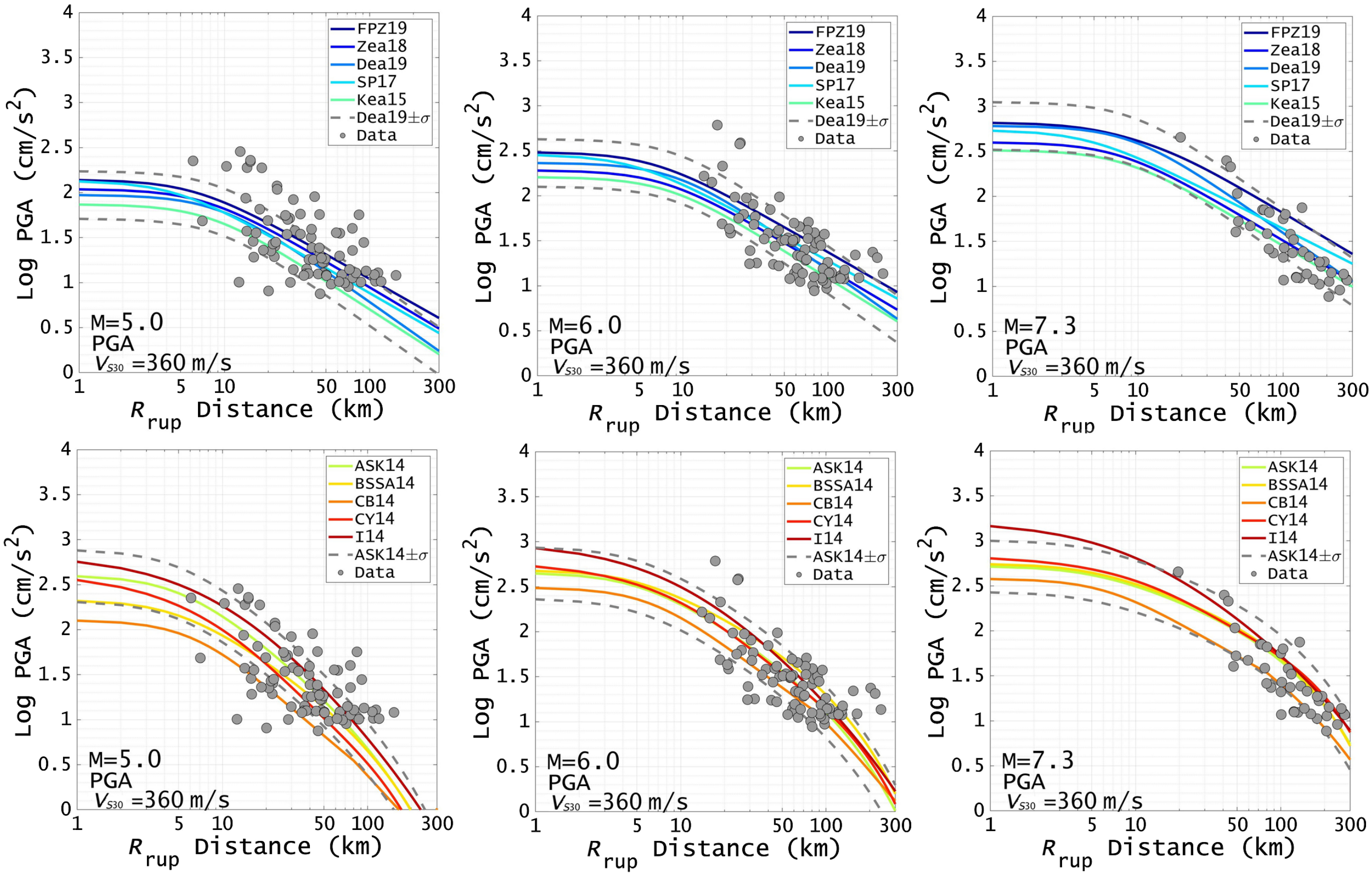

Figure 7. The attenuation of the selected GMMs, local (top row), and NGA-West2 (bottom row), versus distance for PGA. The GMMs are evaluated at $M_w = 5.0$, 6.0, and 7.3 (columns left to right, respectively) for a $V_{S30} = 360$ m/s. Gray circles are the observed strong-motion data. The color version of this figure is available only in the electronic edition.

country. In this regard, the GMMs, as one of the essential elements of any PSHA, should be carefully selected. In this study, we reviewed three data-driven selection methods, EDR, LLH, and DIC, to evaluate the performance of ten empirical GMMs against a recently recorded and independent Iranian strong-motion data set. The data set includes the catastrophic Sarpol-e Zahab earthquake, one of the major destructive events that have occurred in the northwestern ZFTB. On the other hand, because the data set used in this study is limited and only has two earthquakes with $M_w > 6.1$, the variability is a within-event variability, which is expected to be smaller than the variability expected if several events were included.

The results were presented for PGA and 5% damped PSA over the different selected range of periods. The results indicate the best fit between the local GMMs and the observed data. Among the five local models, the GMMs proposed by Darzi *et al.* (2019) and Farajpour *et al.* (2019) fit the recorded data well, particularly at short periods (i.e., PGA and PSA at $T = 0.2$ s). However, at long periods (i.e., $T = 1.0$ and 2.0 s), the models developed by Zafarani *et al.* (2018) and Sedaghati and Pezeshk (2017) are preferable. Moreover, the Kale *et al.* (2015) model shows better performance at long periods (i.e., $T = 1.0$ and 2.0 s) rather than short periods.

The NGA-West2 models are developed based on worldwide shallow crustal earthquakes data and rarely used Iranian strong-motion data. However, the models proposed by Boore *et al.* (2014) and Abrahamson *et al.* (2014) are preferable. We conclude that the NGA-West2 models can be used but not as a backbone model for Iran. They can also be applied with lower weights if a logic-tree format is used. The distribution of the residuals versus distance at all the selected periods exhibits that they overpredict at near-fault distances and underpredict at far-field distances.

As part of this study, we show that in some cases, the EDR and LLH methods favor models with smaller and larger sigma, respectively, regardless of what the true uncertainty is. In contrast, the DIC method that uses Bayesian statistics is shown to optimize the GMM selection for a given region leading to an unbiased assessment of earthquake hazard. The effect of sigma on the performance of these data-driven selection methods is shown and discussed in Kowsari, Halldorsson, Hrafnkelsson, and Jonsson (2019) where several different synthetic data sets were generated and a generic empirical GMM calibrated to it. Furthermore, the DIC method offers an advantage of partially removing the ergodic assumption from the GMM selection

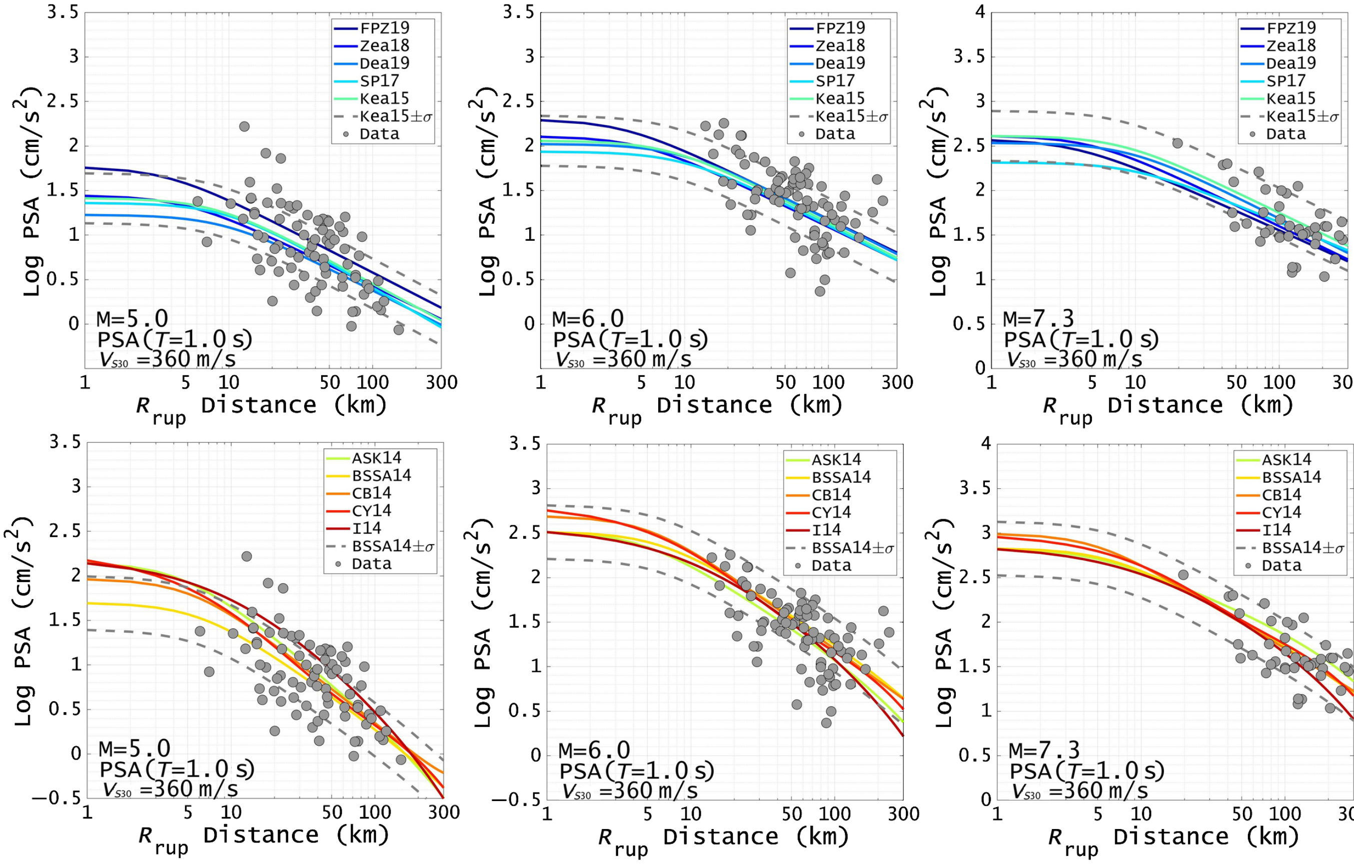

Figure 8. The attenuation of the selected GMMs, local (top row), and NGA-West2 (bottom row), versus distance for PSA at $T = 1.0$ s. The GMMs are evaluated at $M_w = 5.0$, 6.0, and 7.3 (columns left to right, respectively) for a $V_{S30} = 360$ m/s. Gray circles are the observed strong-motion data. The color version of this figure is available only in the electronic edition.

process when the ground-motion recordings are homogeneously distributed in terms of magnitudes and distances. That, in turn, has important implications for the standard PSHA practice.

Finally, the results show that the scores of the data-driven methods for most of the selected GMMs are close to each other. This generally means that those GMMs, with the insignificant difference in their testing score, predict the Iranian strong motions similarly. However, to consider the epistemic uncertainties associated with GMMs through the backbone approach in PSHA, one single GMM (backbone model) is needed. Thus, such insignificant differences between the testing scores matter when the purpose is to rank different GMMs and select the best one as the backbone GMM.

DATA AND RESOURCES

In this study, the strong ground motion records were provided by the Building and Housing Research Center (BHRC) of Iran (http://www.bhrc.ac.ir/, last accessed December 2019). We express our sincere appreciation to the BHRC of Iran that provided strong ground motion records. The flat file of the data set used is provided in the supplemental material to this article. Supplemental plots for residuals corresponding to spectral accelerations (i.e., SA 0.2, 1.0, and 2.0 s), and plot of the attenuation of categories 1 and 2 ground-motion models (GMMs) versus distance and compare them with the observed data for pseudoacceleration spectra (PSA) at $T = 0.2$ and 2.0 s, respectively, are provided in the supplemental material to this article.

ACKNOWLEDGMENTS

The authors thank Fabrice Cotton and two anonymous reviewers for their valuable comments, which led to further improvements in the article. The second and fourth authors would like to acknowledge the Icelandic Centre for Research for funding this research under the Project Grant (Number 196089-051). The authors would like to thank Esmael Farzanegan, Building and Housing Research Center (BHRC), and Mohamad-Reza Ebrahimi to provide data and Generic Mapping Tools (GMT) maps.

REFERENCES

Abrahamson, N. A., W. J. Silva, and R. Kamai (2014). Summary of the ASK14 ground motion relation for active crustal regions, *Earthq. Spectra* **30,** no. 3, 1025–1055.

Atkinson, G. M., and J. Adams (2013). Ground motion prediction equations for application to the 2015 Canadian national seismic hazard maps, *Can. J. Civil Eng.* **40,** no. 10, 988–998.

Atkinson, G. M., and K. Goda (2011). Effects of seismicity models and new ground-motion prediction equations on seismic hazard assessment for four Canadian Cities effects of seismicity models and ground-motion prediction equations on seismic hazard assessment, *Bull. Seismol. Soc. Am.* **101,** no. 1, 176–189.

Atkinson, G. M., J. J. Bommer, and N. A. Abrahamson (2014). Alternative approaches to modeling epistemic uncertainty in ground motions in probabilistic seismic-hazard analysis, *Seismol. Res. Lett.* **85,** no. 6, 1141–1144.

Beauval, C., H. Tasan, A. Laurendeau, E. Delavaud, F. Cotton, P. Guéguen, and N. Kuehn (2012). On the testing of ground-motion prediction equations against small-magnitude data, *Bull. Seismol. Soc. Am.* **102,** no. 5, 1994–2007.

Bommer, J. J., J. Douglas, F. Scherbaum, F. Cotton, H. Bungum, and D. Fäh (2010). On the selection of ground-motion prediction equations for seismic hazard analysis, *Seismol. Res. Lett.* **81,** no. 5, 783–793.

Bommer, J. J., F. Scherbaum, H. Bungum, F. Cotton, F. Sabetta, and N. A. Abrahamson (2005). On the use of logic trees for ground-motion prediction equations in seismic-hazard analysis, *Bull. Seismol. Soc. Am.* **95,** no. 2, 377–389.

Boore, D. M., and J. J. Bommer (2005). Processing of strong-motion accelerograms: Needs, options and consequences, *Soil Dynam. Earthq. Eng.* **25,** no. 2, 93–115.

Boore, D. M., C. D. Stephens, and W. B. Joyner (2002). Comments on baseline correction of digital strong-motion data: Examples from the 1999 Hector Mine, California, earthquake, *Bull. Seismol. Soc. Am.* **92,** no. 4, 1543–1560.

Boore, D. M., J. P. Stewart, E. Seyhan, and G. M. Atkinson (2014). NGA-West2 equations for predicting PGA, PGV, and 5% damped PSA for shallow crustal earthquakes, *Earthq. Spectra* **30,** no. 3, 1057–1085.

Bozorgnia, Y., N. A. Abrahamson, L. A. Atik, T. D. Ancheta, G. M. Atkinson, J. W. Baker, A. Baltay, D.M. Boore, K.W. Campbell, B. S. J. Chiou, *et al.* (2014). NGA-West2 research project, *Earthq. Spectra* **30,** no. 3, 973–987.

Bradley, B. A., M. W. Stirling, G. H. McVerry, and M. Gerstenberger (2012). Consideration and propagation of epistemic uncertainties in New Zealand probabilistic seismic-hazard analysis, *Bull. Seismol. Soc. Am.* **102,** no. 4, 1554–1568.

Campbell, K. W., and Y. Bozorgnia (2014). NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra, *Earthq. Spectra* **30,** no. 3, 1087–1115.

Cao, T., M. D. Petersen, and A. D. Frankel (2005). Model uncertainties of the 2002 update of California seismic hazard maps, *Bull. Seismol. Soc. Am.* **95,** no. 6, 2040–2057.

Chen, K., W. Xu, P. M. Mai, H. Gao, L. Zhang, and X. Ding (2018). The 2017 Mw 7.3 Sarpol Zahāb earthquake, Iran: A compact blind shallow-dipping thrust event in the mountain front fault basement, *Tectonophysics* **747,** 108–114.

Chiou, B. S.-J., and R. R. Youngs (2014). Update of the Chiou and Youngs NGA model for the average horizontal component of peak ground motion and response spectra, *Earthq. Spectra* **30,** 1117–1153.

Cotton, F., F. Scherbaum, J. J. Bommer, and H. Bungum (2006). Criteria for selecting and adjusting ground-motion models for specific target regions: Application to central Europe and rock sites, *J. Seismol.* **10,** no. 2, 137.

Cramer, C. H. (2001). A seismic hazard uncertainty analysis for the New Madrid seismic zone, *Eng. Geol.* **62,** nos. 1/3, 251–266.

Darzi, A., M. R. Zolfaghari, C. Cauzzi, and D. Fäh (2019). 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.* **109,** no. 3, 1041–1057.

Delavaud, E., F. Scherbaum, N. Kuehn, and T. Allen (2012). Testing the global applicability of ground-motion prediction equations for active shallow crustal regions, *Bull. Seismol. Soc. Am.* **102,** no. 2, 707–721.

Delavaud, E., F. Scherbaum, N. Kuehn, and C. Riggelsen (2009). Information-theoretic selection of ground-motion prediction equations for seismic hazard analysis: An applicability study using Californian data, *Bull. Seismol. Soc. Am.* **99,** no. 6, 3248–3263.

DeMets, C., R. G. Gordon, D. F. Argus, and S. Stein (1994). Effect of recent revisions to the geomagnetic reversal time scale on estimates of current plate motions, *Geophys. Res. Lett.* **21,** no. 20, 2191–2194.

Douglas, J. (2018a). Capturing geographically-varying uncertainty in earthquake ground motion models or what we think we know may change, *European Conf. on Earthquake Engineering Thessaloniki, Greece*, Springer, Cham, Switzerland, 153–181.

Douglas, J. (2018b). Calibrating the backbone approach for the development of earthquake ground motion models, *Best Practice in Physics-Based Fault Rupture Models for Seismic Hazard Assessment of Nuclear Installations: Issues and Challenges Towards Full Seismic Risk Analysis (2nd Best PSHANI)*, Cadarache Chateau, France, May 2018.

Engdahl, E. R., J. A. Jackson, S. C. Myers, E. A. Bergman, and K. Priestley (2006). Relocation and assessment of seismicity in the Iran region, *Geophys. J. Int.* **167,** no. 2, 761–778.

Farajpour, Z., S. Pezeshk, and M. Zare (2019). A new empirical ground-motion model for Iran, *Bull. Seismol. Soc. Am.* **109,** no. 2, 732–744.

Farajpour, Z., M. Zare, S. Pezeshk, A. Ansari, and E. Farzanegan (2018). Near-source strong motion database catalog for Iran, *Arabian J. Geosci.* **11,** no. 4, 80.

Farhadi, A., Z. Farajpour, and S. Pezeshk (2019). Assessing predictive capability of ground-motion models for probabilistic seismic hazard in Iran, *Bull. Seismol. Soc. Am.* **109,** no. 5, 2073–2087.

Hintersberger, E., F. Scherbaum, and S. Hainzl (2007). Update of likelihood-based ground-motion model selection for seismic hazard analysis in western central Europe, *Bull. Earthq. Eng.* **5,** no. 1, 1–16.

Idriss, I. M. (2014). An NGA-West2 empirical model for estimating the horizontal spectral values generated by shallow crustal earthquakes, *Earthq. Spectra* **30,** no. 3, 1155–1177.

Kale, Ö., and S. Akkar (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,** no. 2A, 1069–1084.

Kale, Ö., S. Akkar, A. Ansari, and H. Hamzehloo (2015). A ground-motion predictive model for Iran and Turkey for horizontal PGA, PGV, and 5% damped response spectrum: Investigation of possible regional effects, *Bull. Seismol. Soc. Am.* **105,** no. 2A, 963–980.

Karimzadeh, S., B. Feizizadeh, and M. Matsuoka (2019). DEM-based V_{S30} map and terrain surface classification in nationwide scale—A case study in Iran, *ISPRS Int. J. Geo-Inf.* **8,** no. 12, 537.

Khoshnevis, N., R. Taborda, S. Azizzadeh-Roodpish, and C. H. Cramer (2017). Seismic hazard estimation of northern Iran using smoothed seismicity, *J. Seismol.* **21,** no. 4, 941–964.

Kowsari, M., N. Eftekhari, A. Kijko, E. Y. Dadras, H. Ghazi, and E. Shabani (2019). Quantifying seismicity parameter uncertainties and their effects on probabilistic seismic hazard analysis: A case study of Iran, *Pure Appl. Geophys.* **176,** no. 4, 1487–1502.

Kowsari, M., B. Halldorsson, N. Eftekhari, and J. Þ. Snæbjörnsson (2018). Sensitivity analysis of earthquake hazard in Húsavík, North Iceland from variable seismicity and ground motion models, *16th European Conf. on Earthquake Engineering (16ECEE)*, Thessaloniki, Greece, 18–21 June 2018, Paper Number 11759.

Kowsari, M., B. Halldorsson, B. Hrafnkelsson, and S. Jonsson (2019). Selection of earthquake ground motion models using the deviance information criterion, *Soil Dynam. Earthq. Eng.* **117,** 288–299.

Kowsari, M., B. Halldorsson, B. Hrafnkelsson, J. Þ. Snæbjörnsson, and S. Jónsson (2019). Calibration of ground motion models to Icelandic peak ground acceleration data using Bayesian Markov Chain Monte Carlo simulation, *Bull. Earthq. Eng.* **17,** no. 6, 2841–2870.

Kulkarni, R. B., R. R. Youngs, and K. J. Coppersmith (1984). Assessment of confidence intervals for results of seismic hazard analysis, *8th World Conf. on Earthquake Engineering*, 263–270.

Lombardi, A. M., A. Akinci, L. Malagnini, and C. S. Mueller (2005). Uncertainty analysis for seismic hazard in Northern and Central Italy, *Ann. Geophys.* **48,** no. 6, doi: 10.4401/ag-3239.

Mak, S., R. A. Clements, and D. Schorlemmer (2014). Comment on "A new procedure for selecting and ranking ground-motion prediction equations (GMPEs): The Euclidean distance-based ranking (EDR) method" by Özkan Kale and Sinan Akkar, *Bull. Seismol. Soc. Am.* **104,** no. 6, 3139–3140.

Mak, S., R. A. Clements, and D. Schorlemmer (2017). Empirical evaluation of hierarchical ground-motion models: Score uncertainty and model weighting, *Bull. Seismol. Soc. Am.* **107,** no. 2, 949–965.

Mousavi, M., A. Ansari, H. Zafarani, and A. Azarbakht (2012). Selection of ground motion prediction models for seismic hazard analysis in the Zagros region, Iran, *J. Earthq. Eng.* **16,** no. 8, 1184–1207.

Petersen, M. D., B. K. Rastogi, E. S. Schweig, S. C. Harmsen, and J. S. Gomberg (2004). Sensitivity analysis of seismic hazard for the northwestern portion of the state of Gujarat, India, *Tectonophysics* **390,** nos. 1/4, 105–115.

Power, M., B. Chiou, N. Abrahamson, Y. Bozorgnia, T. Shantz, and C. Roblee (2008). An overview of the NGA project, *Earthq. Spectra* **24,** no. 1, 3–21.

Sabetta, F., A. Lucantoni, H. Bungum, and J. J. Bommer (2005). Sensitivity of PSHA results to ground motion prediction relations and logic-tree weights, *Soil Dynam. Earthq. Eng.* **25,** no. 4, 317–329.

Scherbaum, F., F. Cotton, and P. Smit (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,** no. 6, 2164–2185.

Scherbaum, F., E. Delavaud, and C. Riggelsen (2009). Model selection in seismic hazard analysis: An information-theoretic perspective, *Bull. Seismol. Soc. Am.* **99,** no. 6, 3234–3247.

Scherbaum, F., J. Schmedes, and F. Cotton (2004). On the conversion of source-to-site distance measures for extended earthquake source models, *Bull. Seismol. Soc. Am.* **94,** no. 3, 1053–1069.

Sedaghati, F., and S. Pezeshk (2017). Partially nonergodic empirical ground-motion models for predicting horizontal and vertical PGV, PGA, and 5% damped linear acceleration response spectra using data from the Iranian plateau, *Bull. Seismol. Soc. Am.* **107,** no. 2, 934–948.

Shoja–Taheri, J., S. Naserieh, and G. Hadi (2010). A test of the applicability of NGA models to the strong ground-motion data in the Iranian plateau, *J. Earthq. Eng.* **14,** no. 2, 278–292.

Stafford, P. J., F. O. Strasser, and J. J. Bommer (2008). An evaluation of the applicability of the NGA models to ground-motion prediction in the Euro-Mediterranean region, *Bull. Earthq. Eng.* **6,** no. 2, 149–177.

Strasser, F. O., N. A. Abrahamson, and J. J. Bommer (2009). Sigma: Issues, insights, and challenges, *Seismol. Res. Lett.* **80,** no. 1, 40–56.

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

Zafarani, H., L. Luzi, G. Lanzano, and M. R. Soghrat (2018). Empirical equations for the prediction of PGA and pseudo spectral accelerations using Iranian strong-motion data, *J. Seismol.* **22,** no. 1, 263–285.

Manuscript received 30 January 2020
Published online 20 October 2020