Bulletin of the Seismological Society of America, Vol. 108, No. 4, pp. 2265–2277, August 2018, doi: 10.1785/0120170330

Ⓔ

Assessing the Applicability of Ground-Motion Models for Induced Seismicity Application in Central and Eastern North America

by Ali Farhadi, Shahram Pezeshk, and Naeem Khoshnevis

Abstract This study aims to present a relatively short list of interim induced proxy ground-motion models (GMMs) most suitable for induced-seismicity application in central and eastern North America (CENA). Induced proxy GMMs are models not established from datasets strictly made of induced events but can be used to predict ground motions from such events. For this purpose, we test the predictive power of a long list of GMMs against a dataset of induced earthquakes using the popular log-likelihood (LLH) method of Scherbaum *et al.* (2009) and its natural extension, known as the multivariate logarithmic score of Mak *et al.* (2017). Our dataset is a subset of data provided by Rennolet *et al.* (2017) and is composed of 2414 time histories from 384 CENA induced events with hypocentral distances below 50 km and moment magnitudes from 3.5 to 5.8. Candidate GMMs are from two categories, including purely empirical models developed from the Next Generation Attenuation-West2 (NGA-West2) database and indigenous models of CENA. The NGA-West2 database contains a large number of shallow small-to-moderate magnitude events from California that may approximate characteristic features of induced events in CENA. Some of the CENA models have considered near-distance saturation for small-to-moderate magnitude range and/or have explicitly modeled source parameter as a function of focal depth that may make them reasonable induced proxy GMMs.

Some models performed better in certain frequencies than others, and not a single model performed the best over the entire frequency range. Overall, three models including Abrahamson *et al.* (2014), Chiou and Youngs (2014), and Atkinson (2015) GMMs outperformed other models. These models are not specifically established for CENA but are properly modeled for magnitude and depth scaling. In addition, stochastic models favored in the low-seismicity region of CENA appear not to perform better than models developed based on conventional statistical and empirical approaches for induced-seismicity applications. The result of this study can be useful in selecting a suite of appropriate GMMs for performing probabilistic seismic hazard assessment.

Electronic Supplement: Tables describing distinctness computed from log-likelihood scores for all pairs of the candidate models at six spectral periods, plots of residuals versus the magnitude and distance, and visual comparisons between observed ground-motion intensities and models' predictions.

Introduction

Over the last decade, many parts of central and eastern North America (CENA) have hosted dozens of induced events resulting from oil and gas activities. For example, on 6 November 2011, the active disposal wells in the immediate vicinity of Prague, Oklahoma, triggered an M_w 5.7 earthquake that caused modified Mercalli intensities up to VIII and led to damages to local buildings (Ellsworth, 2013; Keranen *et al.*, 2013). In addition, in May 2012, fluid injection in Timpson, Texas, resulted in a relatively smaller earthquake of m_{bLg} 4.1 but considerable peak ground acceleration (PGA) of 0.62g (Frohlich *et al.*, 2014). Most recently, hydraulic fracturing stimulated a sequence of widely felt induced earthquakes with the largest moment magnitude of M_w 4.4 in a small region close to Fox Creek, Alberta (Atkinson *et al.*, 2015). Atkinson *et al.* (2015) assessed the impact of considering potential sources of induced seismicity on the overall

Downloaded from https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/4275358/bssa-2017330.1.pdf
by University of Memphis, 16550
on 13 November 2018

hazard level at Fox Creek by performing separate seismic hazard studies before and after the initiation of induced sequences. They concluded that in low-to-moderate seismicity regions, the hazard from man-made activities could significantly surpass that from tectonic earthquakes, especially in the case of sites close to the location of induced events. Therefore, evaluation of seismic hazard from induced earthquakes is of great importance in low-to-moderate seismicity regions located in CENA.

A crucial element in any seismic hazard study that aims at quantifying hazard from induced events is the ground-motion estimation using a set of suitable ground-motion models (GMMs). Because the attenuation characteristics of induced events is different from those of tectonic earthquakes from which existing GMMs have been developed, their resulting ground motions could be different. Induced events have shallower focus, as well as smaller magnitude, compared to their natural counterparts. At shallower depth with heterogeneous nature, regional variations are more significant. Moreover, the vast majority of existing models have been established from earthquakes with magnitudes larger than M_w 5.0 that are confirmed to be inadequate for prediction of smaller-magnitude range (Bommer *et al.*, 2007; Atkinson and Morrison, 2009; Zafarani and Farhadi, 2017). Bommer *et al.* (2016) posed the need to derive application-specific GMMs for induced-seismicity application. This need could be met in the near future as datasets strictly composed of induced events become accessible in the public literature. On the other hand, Douglas *et al.* (2013) statistically compared induced and natural earthquakes and inferred no discernable difference in attenuation trend among these events for the same magnitude range and similar hypocentral distances. In accordance with Douglas *et al.* (2013), Atkinson (2015; hereafter, A15) used a subset of the Next Generation Attenuation-West2 (NGA-West2) database containing only small-to-moderate magnitude data from California natural earthquakes to develop a new GMM suitable for induced-seismicity application. Atkinson and Assatourians (2017) recently indicated that California tectonic data with depths between 2 and 6 km can approximate induced earthquakes in CENA, due to the opposing effect that depth and tectonic setting have on the stress parameter that scales high-frequency ground-motion amplitudes. They introduced characteristic features of natural GMMs established mainly from California data that make them applicable for estimation of seismic hazard from induced events in CENA. According to Atkinson and Assatourians (2017), existing GMMs, including indigenous models of CENA, may also be desirable induced-proxy GMMs if the modeler appropriately considers the near-distance saturation for small-to-moderate earthquakes and explicitly considers source parameter as a function of the focal depth in their functional form.

Assessing a long list of GMMs that might be suitable for induced-seismicity application was beyond the scope of the study of Atkinson and Assatourians (2017). Accordingly, they did not present a direct comparison of several models against their testing dataset to determine a few suitable models among a large number of candidate ground-motion relations. To date, there is no study related to assessing several models against datasets strictly made of induced events. Gupta *et al.* (2017), for instance, evaluated the applicability of two GMMs including the Atkinson (2015) model as well as the Shahjouei and Pezeshk (2016) relation for induced-seismicity application in central and eastern United States. They concluded that the Atkinson (2015) model is a good fit for hypocentral distances up to 60 km, and that the Shahjouei and Pezeshk (2016) model captures the geometric spreading of ground motions at large distances for both induced and tectonic earthquakes.

There is a lack of comprehensive direct studies accounting for introducing a suite of appropriate models in the existing literature. Therefore, this study tests a long list of GMMs against a dataset of induced earthquakes using the log-likelihood (LLH) method (Scherbaum *et al.*, 2009) and its natural extension, known as the multivariate logarithmic score of Mak *et al.* (2017). To assess the applicability of candidate GMMs for induced-seismicity application, we used a dataset of 2414 time histories from 384 CENA induced earthquakes with hypocentral distances below 50 km and in the 3.5–5.8 moment magnitude range. We used Rennolet *et al.* (2017) as our test data. The number of ground-motion records is reduced for larger spectral ordinates. Candidate GMMs include models developed by the NGA-West2 project (Bozorgnia *et al.*, 2014) conducted by the Pacific Earthquake Engineering Research Center and indigenous models of CENA constrained by regional data. The NGA-West2 database contains a large portion of shallow small-to-moderate magnitude data from California. Some of the CENA models have considered near-distance saturation for the small-to-moderate magnitude range and/or have explicitly modeled source parameter as a function of focal depth.

This study continues by reviewing applied GMMs from the two categories and then presents an introduction to our database. We discuss the goodness-of-fit measures used to evaluate the performance of selected GMMs against testing dataset. Finally, we conclude by providing a relatively short list of appropriate GMMs most suitable for the induced-seismicity application.

Candidate-Induced Proxies GMMs

In this study, we summarize candidate-induced proxy models from two groups of GMMs. The first set of models (group 1) is comprised of purely empirical ones developed mainly from the California data. The second group (group 2) contains models developed or calibrated in accordance with ground motions generated from the CENA data. Table 1 summarizes some characteristic features of candidate models, including their applicability range, their distance metrics, the maximum source-to-site distance that they can cover, which group they belong to, and some additional information.

Downloaded from https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/4275358/bssa-2017330.1.pdf
by University of Memphis, 16550
on 13 November 2018

Table 1
Summary on Applied Ground-Motion Prediction Equations (GMPEs)

GMPE	Abbreviation	Category	Distance Metric	Magnitude Range	Maximum Distance Coverage (km)
Atkinson (2015)	A15	Group 1	R_{hypo}	3–6	40
Abrahamson *et al.* (2014)	ASK14	Group 1	R_{rup}, R_{JB}	3–8.5	300
Boore *et al.* (2014)	BSSA14	Group 1	R_{JB}	3–8.5	400
Campbell and Bozorgnia (2014)	CB14	Group 1	R_{rup}, R_{JB}	3.3–8.5	300
Chiou and Youngs (2014)	CY14	Group 1	R_{rup}, R_{JB}	3.5–8.5	300
Hassani and Atkinson (2015)	HA15	Group 2	R_{JB}	3–8.5	400
Pezeshk *et al.* (2015)*	PZCT15	Group 2	R_{rup}	3–8	1000
Shahjouei and Pezeshk (2016)*	SP16	Group 2	R_{JB}	5–8	1000
Scaled Shahjouei and Pezeshk (2016)*	$SP16_{scaled}$	Group 2	R_{JB}	3–6	200
Yenier and Atkinson (2015)†	YA15	Group 2	R_{rup}	3–8	600

R_{hypo}, hypocentral distance; R_{JB}, closest distance to the surface projection of fault plane; R_{rup}, closest distance to fault plane.

*Predictions for site conditions other than referenced site condition are according to Boore and Campbell (2017).

†Sigma components are according the study of Al Atik (2015).

The first group includes the NGA-West2 GMMs in addition to the model of Atkinson (2015). The Atkinson (2015) model is developed to be applicable for induced seismicity. Atkinson (2015) assumed that the amplitude of motions from induced events is similar to that of tectonic earthquakes with the same magnitudes and hypocentral distances. Although predictions of the Atkinson (2015) model were in good agreement with the induced data of Douglas *et al.* (2013), she mentions the need for performing an independent study to test the applicability of her model for induced events.

The NGA-West2 project (Bozorgnia *et al.*, 2014) database comprises a relatively large portion of small-to-moderate magnitude data from tectonic earthquakes that might be similar to induced events in terms of key features such as the magnitude and the distance scaling, which control ground-motion amplitudes. One of the goals of the most recent NGA-West2 model developers was to update their own models and make their models applicable to smaller magnitude data, taking advantage of the rich database provided by Ancheta *et al.* (2014). By excluding the model of Idriss (2014) that is valid only for sites with V_{S30} values larger than 400 m/s, we considered four NGA models, including Abrahamson *et al.* (2014; hereafter, ASK14), Boore *et al.* (2014; hereafter, BSSA14), Campbell and Bozorgnia (2014; hereafter, CB14), and Chiou and Youngs (2014; hereafter, CY14).

The second group of the applied GMMs considered for this study includes models constrained based on events that occurred in CENA. Here, we use the referenced empirical model of Hassani and Atkinson (2015; hereafter, HA15) as the first representative of the models calibrated mainly from the CENA data. HA15 performed residual analysis to adjust the reference model of BSSA14 to match the observed data in eastern North America.

We also evaluated the applicability of the hybrid empirical model of Shahjouei and Pezeshk (2016; hereafter, SP16) against the testing database. Most recently, Gupta *et al.* (2017) scaled the SP16 model in accordance with the prediction of the A15 model at short hypocentral distances and the distance attenuation of the SP16 model. They proposed the $SP16_{scaled}$ model as more appropriate GMMs for both the ground-motion amplitudes and attenuation to 200 km than the A15 and SP16 GMMs. In addition to the SP16 model, we considered another hybrid empirical model introduced in Pezeshk *et al.* (2015; hereafter, PZCT15). PZCT15 presented an update to the model of Pezeshk *et al.* (2011) by considering the five NGA-West2 GMMs and the latest information on CENA and western North America seismological parameters. It is important to mention that the three hybrid GMMs provide their prediction for their definition of referenced site conditions. Therefore, we implemented the method of Boore and Campbell (2017) to modify predictions of SP16, $SP16_{scaled}$, and PZCT15 for site conditions other than their reference site conditions.

Yenier and Atkinson (2015; hereafter, YA15) took the advantages of both hybrid and referenced empirical methods and proposed a new simulation-based generic model that can be adjusted to any region in accordance with empirical data of the target region. YA15 used the NGA-East database to calibrate their generic model to be suitable for use in CENA. The YA15 model does not present sigma values. Therefore, for this model we applied the total standard deviation (σ), between-event standard deviation (τ), and within-event standard deviation (ϕ), in accordance with proposed values of Al Atik (2015) for CENA.

Testing Dataset

Engineers are generally interested in induced earthquakes with hypocentral distances less than 50 km, in which a shallow-depth small-to-moderate event can be a damaging event to the local structures. As a result, we focused our evaluation on the magnitude–distance range of most concern to hazard estimation from induced events, M_w 3.5–6 at distances to 50 km (Atkinson and Assatourians, 2017).

Downloaded from https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/4275358/bssa-2017330.1.pdf
by University of Memphis, 16550
on 13 November 2018

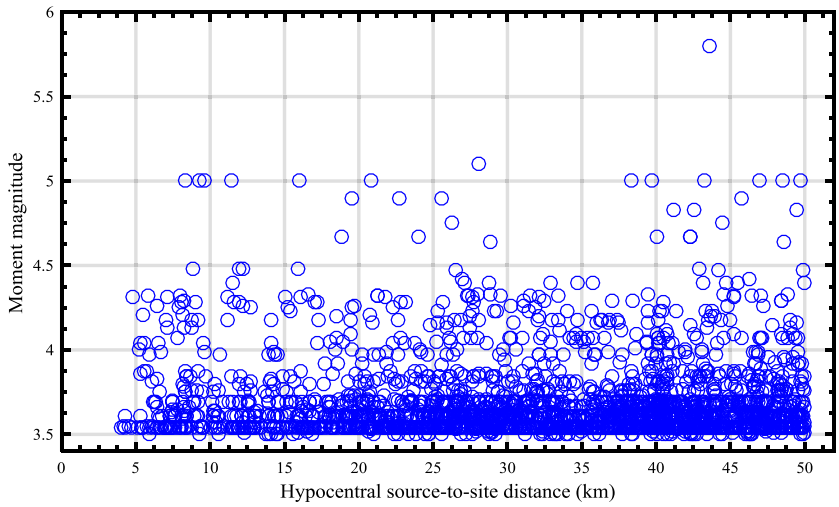

Figure 1. Magnitude–distance distribution of the testing dataset. The color version of this figure is available only in the electronic edition.

The evaluation database is a subset of data compiled by Rennolet *et al.* (2017). Rennolet *et al.* (2017) provided a database of orientation-independent ground-motion intensity measures (RotD50 and RotD100) from 3800 induced events with moment magnitude $M_w \geq 3$ that occurred in Oklahoma and Kansas from January 2009 to December 2016. It is worth mentioning that all candidate models provide their predictions for the RotD50 parameter and are compatible with the evaluation dataset. To obtain an evaluation database independent of candidate models, we removed those events included in the NGA-East flatfile from the database of Rennolet *et al.* (2017). Therefore, our evaluation dataset is independent of all candidate models, and results of the goodness-of-fit measures will compare the predictive power of the evaluated models rather than their explanatory power (Mak, 2017).

Overall, our evaluation database is comprised of 2414 time histories from 384 CENA induced earthquakes with hypocentral distances below 50 km and moment magnitudes ranging from 3.5 to 5.8. The number of earthquakes and usable records are shown in Table 2. As is clear from Table 2, the size of the evaluation database is reduced at higher spectral periods. Figure 1 displays the distribution of the magnitude data versus associated hypocentral distances for PGA. Recording stations and earthquake locations are illustrated in Figure 2. This figure includes 384 earthquakes recorded by 132 stations. Rennolet *et al.* (2017) provided V_{S30} values determined mainly from the topographic slope method of Wald and Allen (2007). This method does not measure V_{S30} values but provides interim proxies for seismic site conditions. Nonetheless, for stations with available estimated V_{S30} values from various proxy-based methods, we averaged the associated values to have a more reasonable estimate of the local site effects. V_{S30} values estimated from proxy-based methods other than the slope method of Wald and Allen (2007) are obtained from the studies of Hosseini *et al.* (2016), Parker *et al.* (2017), and Zalachoris *et al.* (2017). Clearly, for sites with direct measurements of shear-wave velocities, measured values are considered instead of the average V_{S30} values obtained from proxy-based methods.

Table 2
Number of Earthquakes and Records Used for Evaluating Ground-Motion Models at Six Representative Periods

Period	N_{events}	$N_{records}$
PGA	384	2414
$T = 0.2$ s	384	2413
$T = 0.5$ s	382	2274
$T = 1.0$ s	376	1892
$T = 2.0$ s	340	1370
$T = 3.0$ s	264	892

PGA, peak ground acceleration.

An important challenge faced when the NGA models are tested against a database is the estimation of input parameters used for each candidate model. We tried to handle this challenge as much as possible by estimating unknown values using references by the NGA GMMs developers (e.g., Kaklamanos *et al.*, 2011). In addition, the newer version of these models let the users assume some poorly constrained input parameters as unknown values. This feature reduces the possibility of coming up with biased predictions. Moreover, all candidate models require extended source-to-site distances as input parameters. In this regard, we assumed the Joyner–Boore distance to be equal to the epicentral distance ($R_{JB} = R_{epi}$) and the closest distance to the fault-ruptured area as the hypocentral distance ($R_{rup} = R_{hyp}$). This is due to a lack of information about the geometry of a fault's ruptured plane for small-to-moderate magnitude data. These assumptions are physically consistent with the small ruptured area for this magnitude range, and we made them in accordance with previous experiences and arguments raised by Ambraseys *et al.* (2005), Bindi *et al.* (2006), Goda and Atkinson (2014), Farhadi and Mousavi (2016), and Tavakoli *et al.* (2018).

Tests of Goodness-of-Fit Measures

In this study, we used two approaches to examine relative performance of the candidate models. We implemented the popular LLH method of Scherbaum *et al.* (2009) and its natural extension known as the multivariate logarithmic score (mvLogS) of Mak *et al.* (2017). For the LLH method, one needs to compute the probability density function (PDF) of a given observation by assuming a normal distribution for the logarithmic predictions of each candidate GMM. This

Downloaded from https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/4275358/bssa-2017330.1.pdf
by University of Memphis, 16550
on 13 November 2018

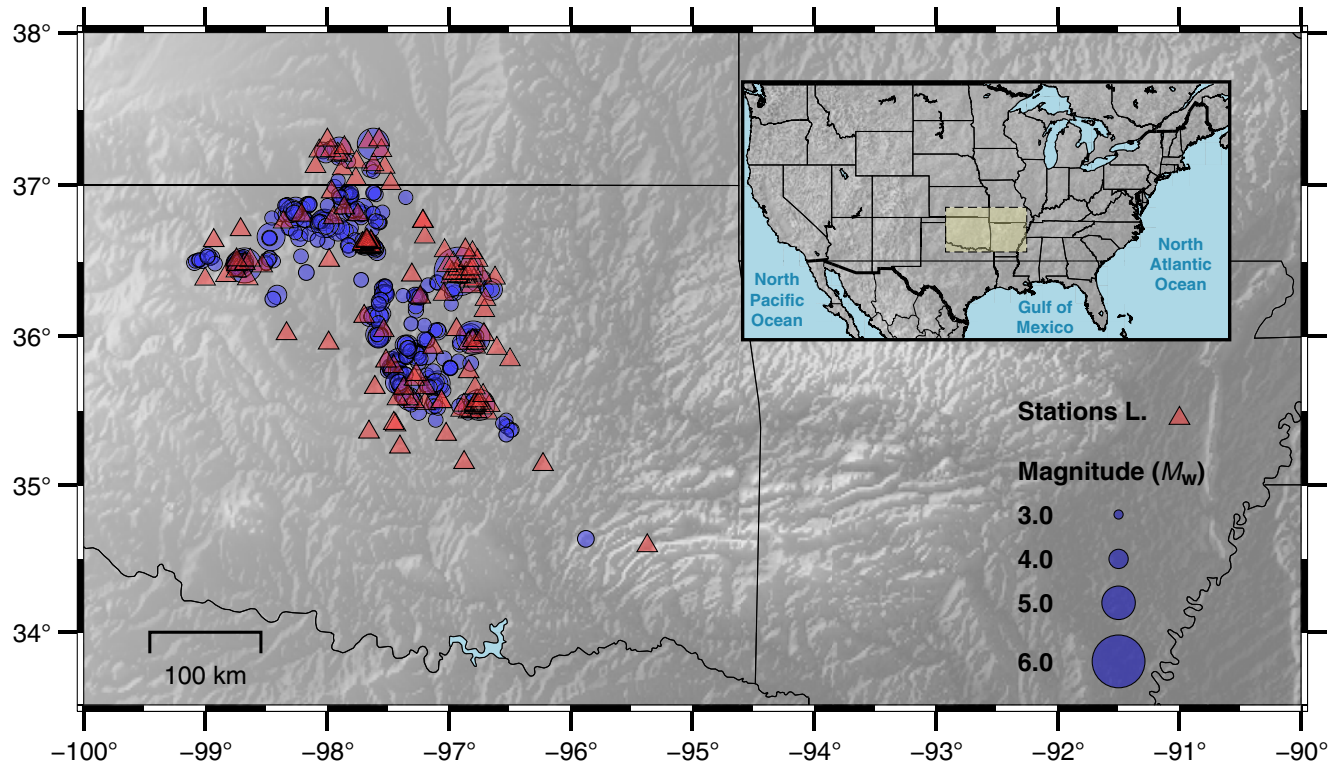

Figure 2. Earthquakes and stations considered for the test data. (Inset) Geographical location of the study region within the United States. The color version of this figure is available only in the electronic edition.

normal distribution is characterized by the mean and the standard deviation equal to the logarithmic median prediction and the standard deviation in the log unit of the model. In other words, for an individual observation x_i, we first compute the LLH value from $\log_2(g(x_i))$, in which $g(x_i)$ is the PDF of the candidate GMM. Then, the LLH score is the average of the LLH values computed for all observed data from equation (1), with N representing the total number of observations:

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

The original LLH method is insensitive to sigma partition and ignores the correlation structure of GMMs in the evaluation process. In other words, this method is unable to use all information provided by the modelers of recently developed GMMs that partition the total sigma into various components. The mvLogS, however, exploits all information provided for sigma components and addresses the correlation structure of hierarchical GMMs in evaluating a model's relative performances. In addition, this approach is less sensitive to unbalanced data and is less likely to be biased toward events with larger ground-motion records. The mvLogS approach is a natural extension of the LLH method and takes the effect of correlation structure into account for evaluating the model's performance. Scores are computed from the following equation:

$$\mathrm{mvLogS} = [N\ln(2\pi) + \ln|\mathbf{V}| + (\mathbf{q}-\mathbf{p})'\mathbf{V}^{-1}(\mathbf{q}-\mathbf{p})]/2, \tag{2}$$

in which N represents the total number of observations; $\mathbf{p}$ and $\mathbf{q}$ are the vectors of logarithmic predictions and logarithms of observed ground motions, in turn; and $|\mathbf{V}|$ and $\mathbf{V}^{-1}$ are the determinant and the inverse of the covariance matrix, respectively (see Mak *et al.*, 2017, for further information).

Both methods generalize a model's relative performance by presenting a score for each candidate model. Scores are random numeric values and may change if one changes the evaluation dataset. Therefore, it may not be reasonable to rank the candidate models merely based on the absolute values of their scores. To assess score variability, we used a cluster bootstrap technique at the event level to resample 400 datasets from the evaluation database. Cluster bootstrap can be generalized into a single value, referred to as the distinctness index (DI), that shows if the two models are truly different, given the score variability. DI ranges from −1.0 to 1.0, and a positive value of DI indicates that the model scores better more often than another one, given the variability of the evaluation data. In this study, we used DI values to rank candidate models, instead of ranking them based on the absolute value of their final score. A model having all positive DIs is the best model and more often than not scores better than rest of the GMMs. However, the second-best model should have a single negative DI with respect to the best model. One may compute the DI from the following equation:

$$\mathrm{DI}_{ij} = \frac{1}{N_{\mathrm{bs}}}\sum_{k}^{N_{\mathrm{bs}}}\tilde{1}(s_i^{(k)}, s_j^{(k)})$$

Downloaded from https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/4275358/bssa-2017330.1.pdf
by University of Memphis, 16550
on 13 November 2018

Model i / Model j	ASK14	A15	BSSA14	CB14	CY14	PZCT15	HA15	SP16	$SP16_{scaled}$	YA15	mvLogS	Rank
ASK14		–1.00	0.13	–1.00	0.93	1.00	1.00	1.00	0.98	1.00	2896	4
A15	1.00		1.00	0.95	1.00	1.00	1.00	1.00	1.00	1.00	2772	1
BSSA14	0.13	–1.00		–1.00	0.99	1.00	1.00	1.00	0.99	1.00	2893	3
CB14	1.00	–0.95	1.00		1.00	1.00	1.00	1.00	1.00	1.00	2820	2
CY14	–0.93	–1.00	–0.99	–1.00		1.00	1.00	1.00	0.88	1.00	2934	5
PZCT15	–1.00	–1.00	–1.00	–1.00	–1.00		–1.00	1.00	–1.00	–1.00	3931	9
HA15	–1.00	–1.00	–1.00	–1.00	–1.00	1.00		1.00	–1.00	–0.97	3130	8
SP16	–1.00	–1.00	–1.00	–1.00	–1.00	–1.00	–1.00		–1.00	–1.00	4429	10
$SP16_{scaled}$	–0.98	–1.00	–0.99	–1.00	–0.88	1.00	1.00	1.00		0.92	2986	6
YA15	–1.00	–1.00	–1.00	–1.00	–1.00	1.00	0.97	1.00	–0.92		3070	7

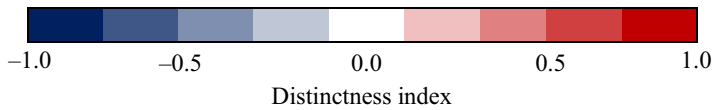

Figure 3. Distinctness table for the candidate models for peak ground acceleration (PGA). Distinctness index of each pairwise comparison (based on 400 bootstrap samples) given in the intersecting box of a model pair. The whole dataset is considered to compute multivariate logarithmic scores (mvLogS) in the second last column. ASK14, Abrahamson *et al.* (2014); A15, Atkinson (2015); BSSA14, Boore *et al.* (2014); CB14, Campbell and Bozorgnia (2014); CY14, Chiou and Youngs (2014); PZCT15, Pezeshk *et al.* (2015); HA15, Hassani and Atkinson (2015); SP16, Shahjouei and Pezeshk (2016); YA15, Yenier and Atkinson (2015). The color version of this figure is available only in the electronic edition.

$$\tilde{1}(s_i^{(k)}, s_j^{(k)}) = \begin{cases} 1 & \text{when } s_i^{(k)} < s_j^{(k)} \\ -1 & \text{when } s_i^{(k)} > s_j^{(k)} \\ 0 & \text{when } s_i^{(k)} = s_j^{(k)} \end{cases}, \tag{3}$$

in which DI_{ij} is the DI of model i with respect to model j; N_{bs} represents the number of bootstrap samples; and $s_i^{(k)}$ is the score of model i for the kth bootstrap sample. A model with better performance has a smaller score. $\tilde{1}(.)$ is a modified indicator function.

In the Results section, we present the results according to the LLH method as well as the mvLogS and provide several plots of residuals against magnitude and distance to confirm results.

Results

We present the results for six representative spectral periods including PGA, 0.2, 0.5, 1.0, 2.0, and 3.0 s. Representative periods are selected in a way to compare candidate models over a wide frequency range from low to high frequencies. We ranked models at each spectral period based on distinctness tables taken from DIs of all model pairs. For PGA, we added plots of residuals and we made visual comparisons between observed ground-motion intensities and predicted values to confirm the results. For other spectral periods, however, plots of visual comparisons among candidate models are moved to the Ⓔ electronic supplement to this article (see Ⓔ Figs. S1–S20).

We begin presenting the details of our evaluations by describing the results for the PGA. Figure 3 presents the distinctness table according to the mvLogS for the candidate GMMs at PGA. According to this figure, models developed from the NGA-West2 database outperform CENA regional GMMs. The A15 model is the best model and scores better than rest of the GMMs. CB14 is the second-best model at the PGA and outperforms the rest of the candidate models except A15 in all bootstrap samples. As a result, A15 and CB14 can be considered as models that are truly different from the others for PGA. The DI of 0.13 for BSSA14 with respect to ASK14 shows that the former outperforms the latter in only 56% of the bootstrap samples. To confirm the results for the PGA, plots of normalized residuals (Z_0) are provided in Figures 4 and 5 with respect to both magnitude and distance. Normalized residual can be computed from the following equation:

$$Z_0 = \frac{\ln(SA_{obs}) - \ln(SA_{pre})}{\sigma_{GMM}}, \tag{4}$$

in which SA_{obs} and SA_{pre} represent the observed and predicted spectral accelerations response spectra in a specified period, and σ_{GMM} represents the total standard deviation of the GMM. Figure 4 represents models of group 1, and Figure 5 shows the residual trend for indigenous models of CENA. Comparing these two figures, group 1 models generally outperform group 2 models. In the details, one could see a pronounced negative residual trend nearly at all magnitudes and hypocentral distances for three models with

Downloaded from https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/4275358/bssa-2017330.1.pdf
by University of Memphis, 16550
on 13 November 2018

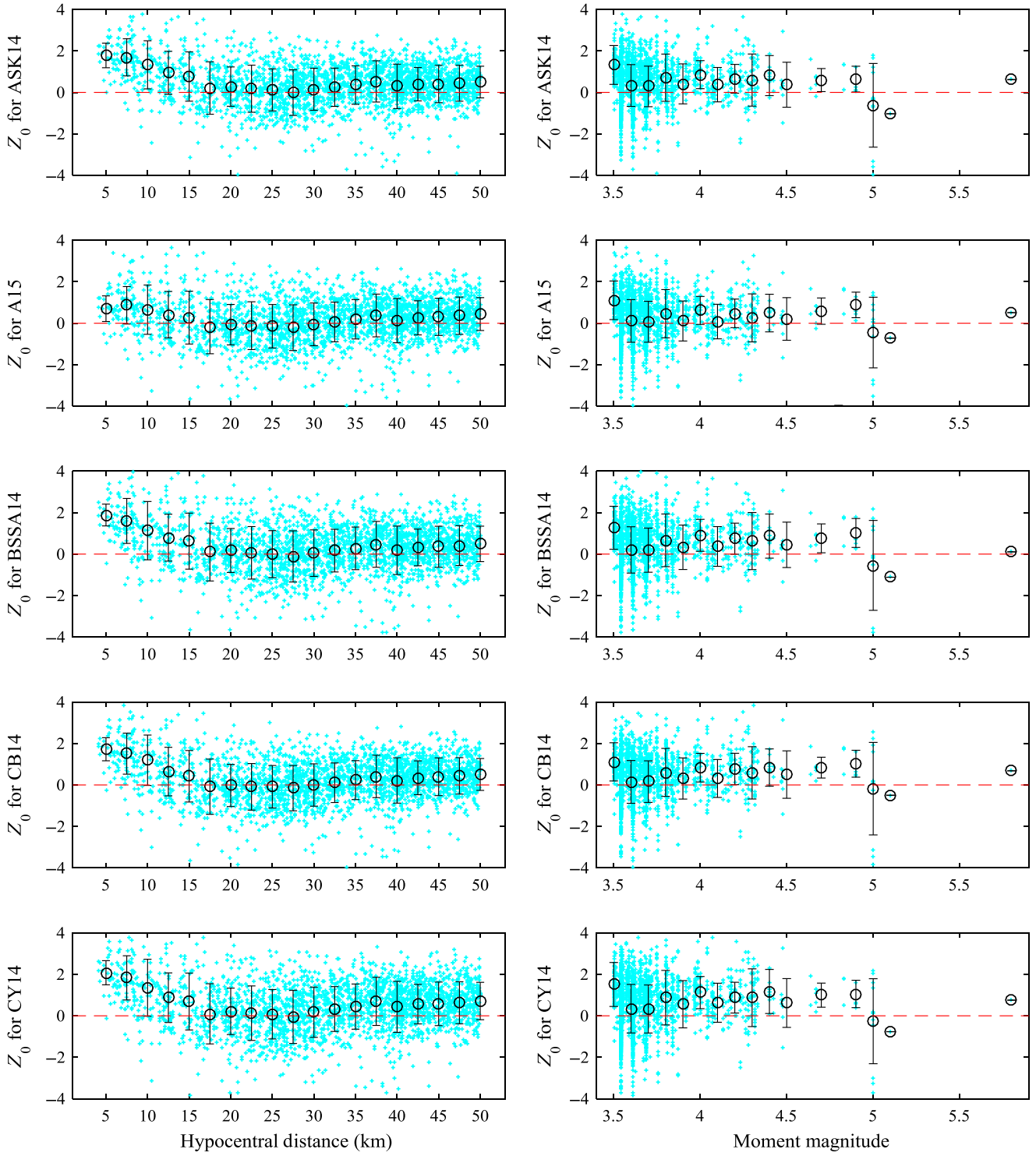

Figure 4. The distribution of the total residuals (Z_0) versus magnitude and hypocentral distance for PGA for empirical models developed from the Next Generation Attenuation-West2 (NGA-West2) database. The mean and the standard deviation of error bars are calculated using distance bins of ± 2.5 km and magnitude bins of ± 0.10 centered at the marker. The color version of this figure is available only in the electronic edition.

the poorest performance at PGA. For more clarification, we provided two more figures to compare observed data with model predictions. Figures 6 and 7 show the comparison of normalized residuals with the standard normal distribution for all candidate GMMs. These two figures confirm the findings of the distinctness table according to the mvLogS goodness-of-fit measure and indicate the poorest performance for the three mentioned models. The difference among models'

Downloaded from https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/4275358/bssa-2017330.1.pdf
by University of Memphis, 16550
on 13 November 2018

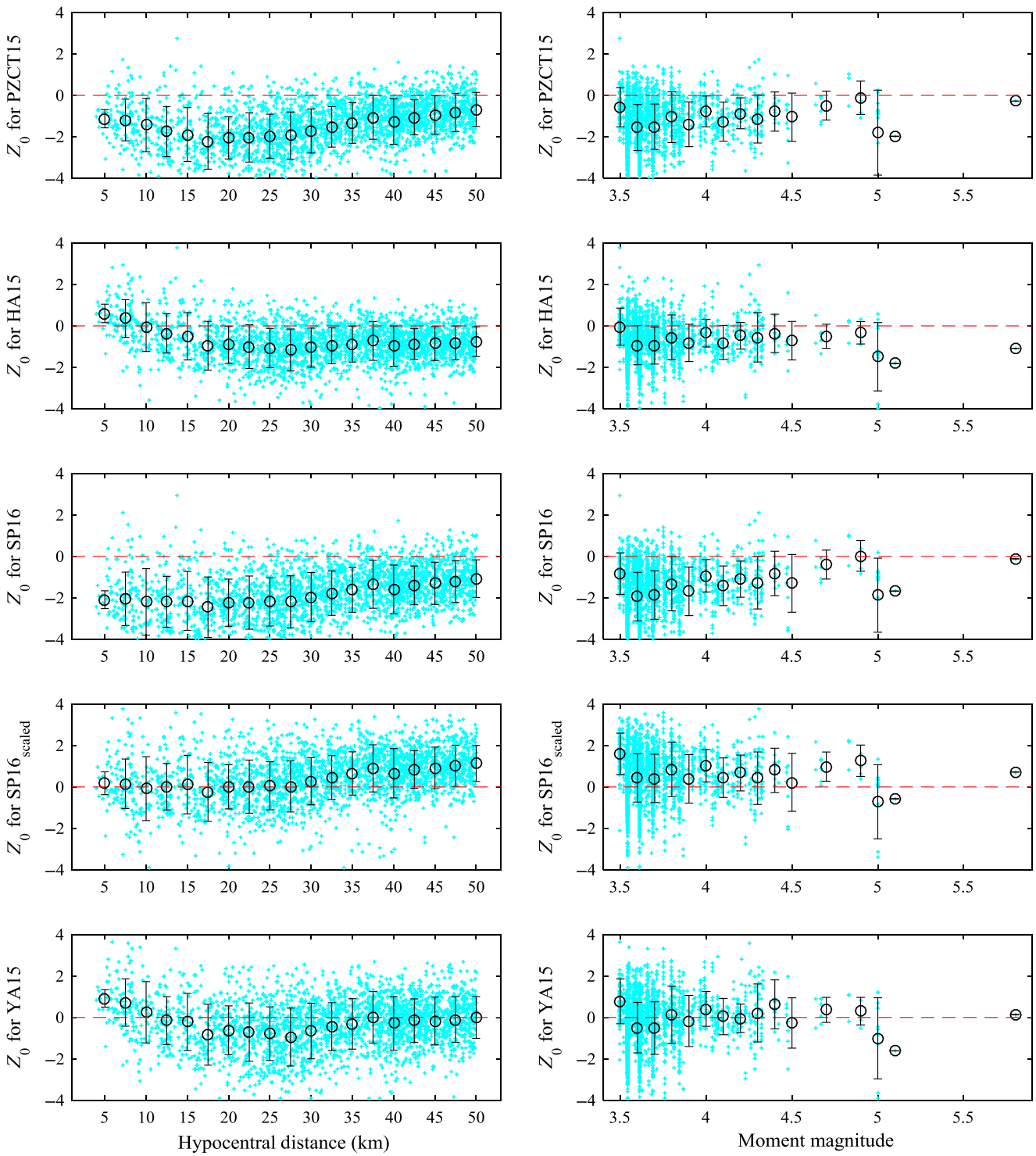

Figure 5. The distribution of the total residuals (Z_0) versus magnitude and hypocentral distance for PGA for indigenous models of central and eastern North America (CENA). The mean and the standard deviation of error bars are calculated using distance bins of ± 2.5 km and magnitude bins of ± 0.10 centered at the marker. The color version of this figure is available only in the electronic edition.

scores is significant for PGA. At PGA, PZCT15 and SP16 models have scores that are significantly larger than scores of other models. The reason for this is a set of calculations performed to modify the model's predictions for site conditions other than their referenced site condition. Modification factor is very large for PGA compared to other spectral periods. This resulted in larger predictions for these two models and overpredicting the observed ground motions.

Ⓔ Figures S21–S25 represent the distinctness tables based on the DIs computed from mvLogS for the five

Downloaded from https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/4275358/bssa-2017330.1.pdf
by University of Memphis, 16550
on 13 November 2018

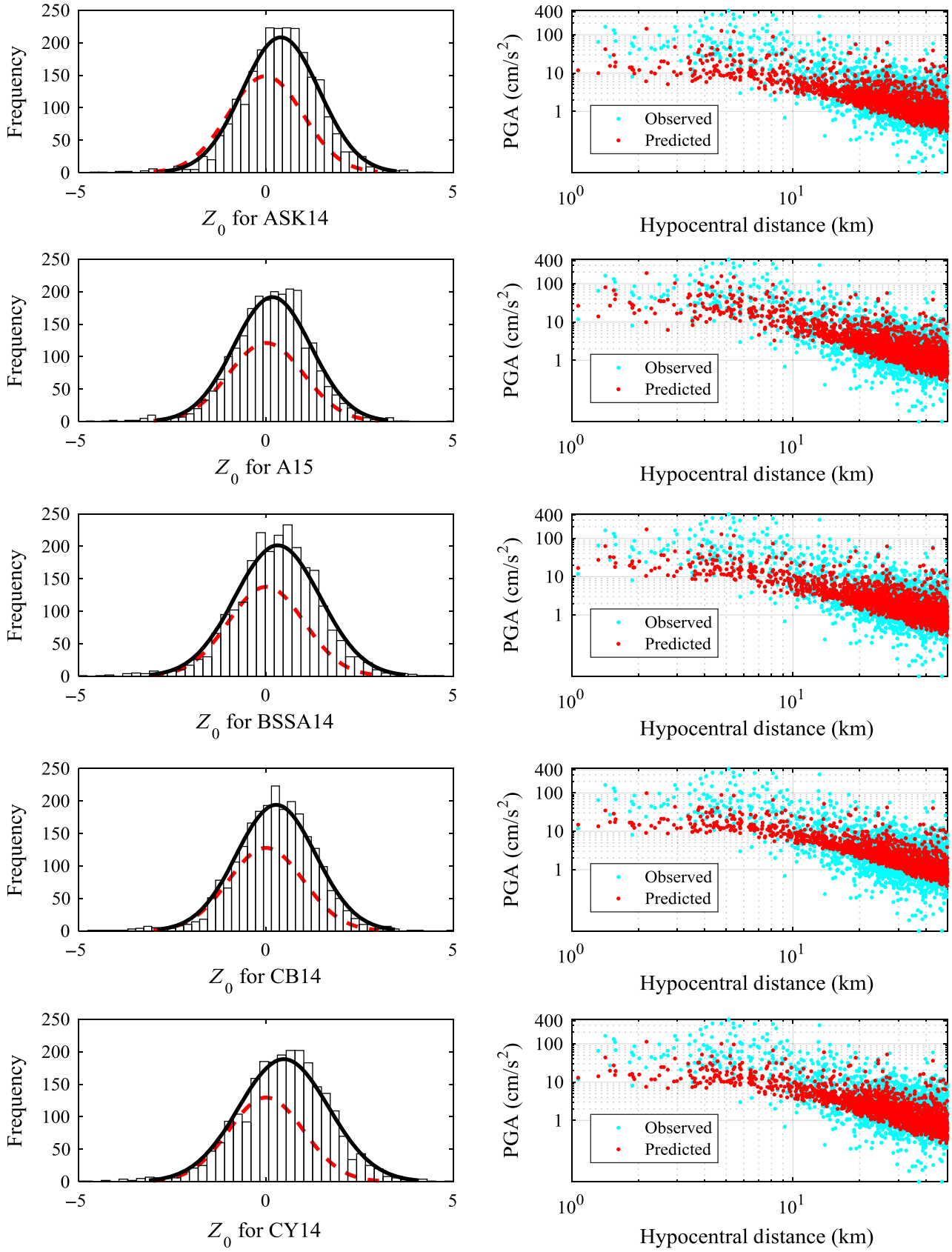

Figure 6. Comparison of normalized residuals (Z_0) with the standard normal distribution (dashed line) as well as a visual comparison of a model's predictions with observed data for empirical models established from the NGA-West2 database for PGA. The color version of this figure is available only in the electronic edition.

Downloaded from https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/4275358/bssa-2017330.1.pdf
by University of Memphis, 16550
on 13 November 2018

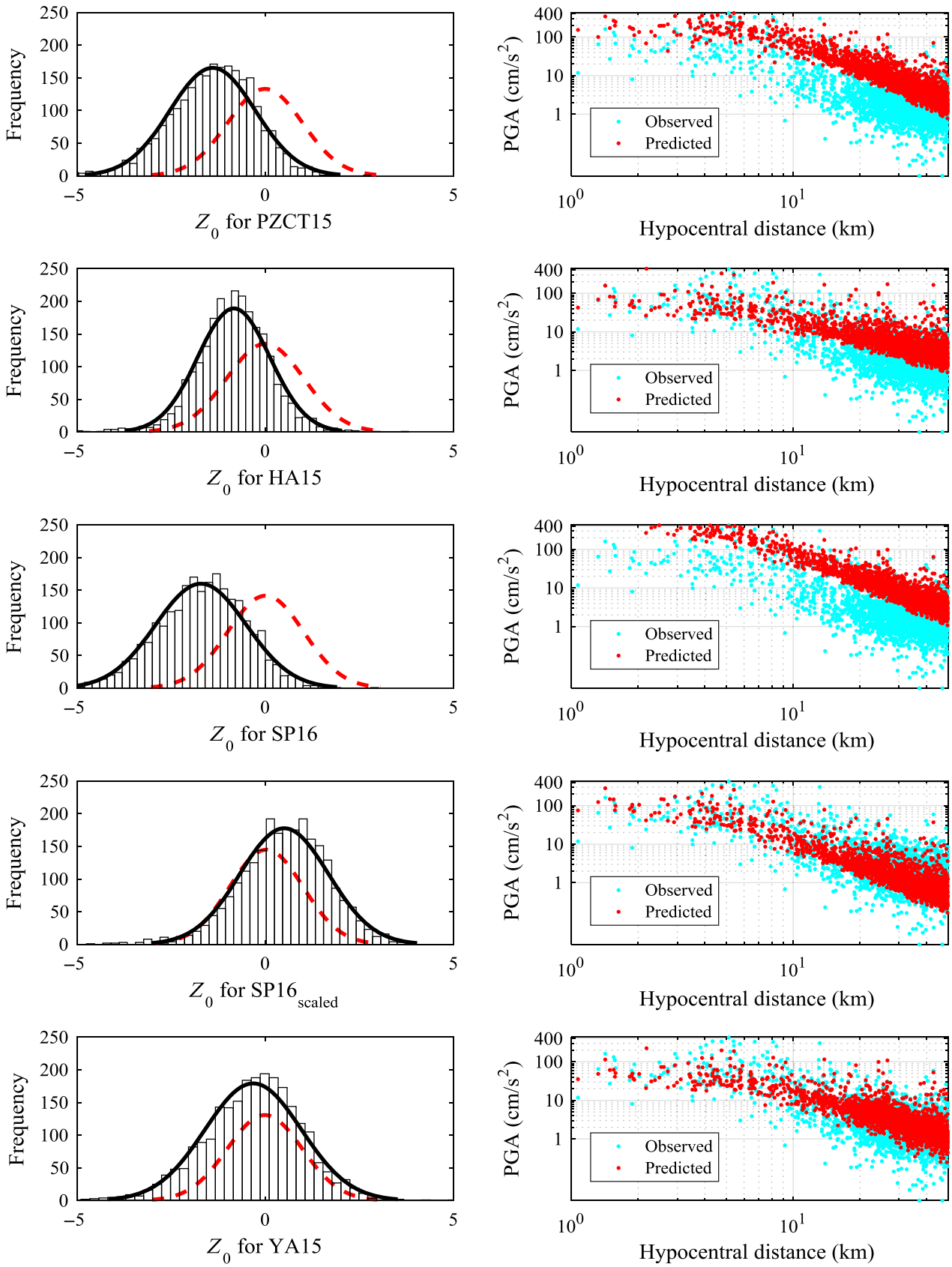

Figure 7. Comparison of normalized residuals (Z_0) with the standard normal distribution (dashed line) as well as visual comparison of a model's predictions with observed data for indigenous models of CENA for PGA. The color version of this figure is available only in the electronic edition.

Downloaded from https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/4275358/bssa-2017330.1.pdf
by University of Memphis, 16550
on 13 November 2018

Table 3
Models' Ranking According to the Multivariate Logarithmic Scores (**mvLogS**) at All Representative Periods

	PGA		$T = 0.2$ s		$T = 0.5$ s		$T = 1.0$ s		$T = 2.0$ s		$T = 3.0$ s	
Model	mvLogS	Rank	mvLogS	Rank	mvLogS	Rank	mvLogS	Rank	mvLogS	Rank	mvLogS	Rank
ASK14	2896	4	3017	6	3066	2	2705	1	2070	5	1414	1
A15	2772	1	2915	1	3071	3	2755	3	2023	2	1441	4
BSSA14	2893	3	2997	4	3212	7	2877	7	2146	7	1511	7
CB14	2820	2	3015	5	3316	9	2949	9	2188	9	1503	6
CY14	2934	5	2948	2	3034	1	2729	2	2043	4	1436	3
PZCT15	3931	9	3051	8	3075	4	2848	5	2159	8	1563	8
HA15	3130	8	3040	7	3225	8	3069	10	2227	10	1578	9
SP16	4429	10	3801	10	3593	10	2899	8	2015	1	1605	10
$SP16_{scaled}$	2986	6	2974	3	3166	6	2847	6	2042	3	1422	2
YA15	3070	7	3063	9	3164	5	2828	4	2069	6	1478	5

Table 4
Models' Ranking According to the Log-Likelihood (LLH) Method at All Representative Periods

	PGA		$T = 0.2$ s		$T = 0.5$ s		$T = 1.0$ s		$T = 2.0$ s		$T = 3.0$ s	
Model	LLH	Rank	LLH	Rank	LLH	Rank	LLH	Rank	LLH	Rank	LLH	Rank
ASK14	2.03	4	2.01	5	2.34	3	2.52	3	2.47	7	2.29	1
A15	1.88	1	1.96	1	2.37	4	2.52	2	2.32	1	2.41	3
BSSA14	2.00	3	2.00	4	2.47	7	2.59	5	2.41	3	2.50	6
CB14	1.93	2	2.01	6	2.72	9	2.90	9	2.57	8	2.35	2
CY14	2.15	6	1.98	3	2.27	2	2.47	1	2.37	2	2.44	4
PZCT15	3.41	9	2.01	7	2.20	1	2.53	4	2.45	4	2.69	8
HA15	2.41	8	2.09	8	2.50	8	2.91	10	2.72	10	2.69	7
SP16	4.13	10	3.13	10	3.08	10	2.74	8	2.46	6	3.24	10
$SP16_{scaled}$	2.15	5	1.97	2	2.42	5	2.66	6	2.46	5	2.48	5
YA15	2.23	7	2.12	9	2.45	6	2.68	7	2.64	9	2.70	9

remaining representative periods. Instead of providing the distinctness tables for individual periods in the main article, we present ranking results based on the multivariate logarithmic scores in Table 3 for all spectral periods. According to Table 3, for the short period of 0.2 s, results are relatively similar to those obtained for PGA, and in general the models of group 1 outperform CENA models. However, three GMMs including CB14, CY14, and a scaled version of SP16 experienced a significant change in their rankings compared to PGA. The last two models with significant improvements perform well and are now the second- and the third-best models, respectively. CB14 and ASK14 show similar performances and experience lower rankings compared to PGA.

By increasing the period from PGA to 0.5 s, we yet observed no significant improvement in the model's ranking for the CENA models, except for the PZCT15 model. This model is now among the top five models. CB14 and BSSA14 are no longer among the most appropriate models, but ASK14 preserved its ranking compared to PGA. Similar to what we observed at 0.2 s, CY14 shows significant improvement at 0.5 s compared to PGA.

Models' rankings at the median period of 1.0 s are relatively similar to those obtained for 0.5 s. The three best models remained the same, and ASK14 scored better than CY14.

At longer periods of 2.0 and 3.0 s, we observed that a larger number of DIs are close to zero (see Ⓔ Figs. S24 and S25), which may indicate more similarity in models' performances compared to the shorter periods. A valid interpretation for the DIs being close to zero is that the models' performances are less stable, due to the reduced sample size for longer periods. Lower stability is linked with the higher fluctuation in the models' relative performances among bootstrap samples, resulting in a DI closer to zero. At longer periods, rankings of CENA models show improvements compared to shorter periods. For example, at period of 2.0 s, SP16 is the best GMM. At 3.0 s, $SP16_{scaled}$ and ASK14 can be considered as the best models. Similar to other periods, A15 and CY14 are still among the five best models.

We continue this section by presenting the results based on the LLH approach. Final rankings based on the distinctness tables computed from the LLH scores and the absolute values of the LLH scores are similar for all representative periods. Therefore, we provided a single table to represent the results according to the LLH approach for all spectral periods and avoided presenting distinctness tables for individual periods. Distinctness tables according to the LLH method can be found in Ⓔ Figures S26–S31. Table 4 shows models' final ranking based on the DIs computed from the

Downloaded from https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/4275358/bssa-2017330.1.pdf
by University of Memphis, 16550
on 13 November 2018

LLH scores. According to this table, results of the LLH method do not show significant variations from the results based on the mvLogS up to 0.2 s. At PGA and 0.2 s, the best and the worst models are the same for both methods, and models' ranking have changed by only one unit. By increasing the spectral period, however, results of the two goodness-of-fit measures are significantly different for some GMMs. Models' similar performances at larger periods could be one reason for this. In addition, we have nearly equal LLH scores for few models at the same period. For example, at a period of 2.0 s, we obtained nearly equal LLH scores of 2.41, 2.45, 2.46, and 2.46 for BSSA14, PZCT15, SP16, and a scaled version of SP16. Another valid reason is that the model's performances are less stable, due to the reduced sample size for longer periods.

Summary and Conclusions

We assessed the applicability of 10 GMMs for induced-seismicity application in CENA using a database of induced earthquakes. Candidate GMMs were obtained from either California or CENA data. Our database is comprised of 384 induced events with 2414 ground-motion records. The evaluation database covers magnitudes from 3.5 to 5.8 and hypocentral distances up to 50 km. We evaluated candidate models against the evaluation database using two methods of ranking, including the popular LLH method of Scherbaum *et al.* (2009) and the multivariate logarithmic score of Mak *et al.* (2017). We ranked the candidate models at six representative spectral periods, including PGA, 0.2, 0.5, 1.0, 2.0, and 3.0 s, based on the distinctness tables taken from the DIs of all model pairs. Models' performances are less stable for larger periods than shorter periods, due to the reduced sample size for longer periods. Models of group 1 outperformed group 2 models for the majority of the representative periods. According to multivariate logarithmic scores, some of the CENA models, including SP16 and its scaled version, perform well for larger periods of 2.0 and 3.0 s, respectively, and are comparable with group 1 models. As a result, models' ranking changes over the frequency range, and various models from both groups may be considered as the most suitable induced proxy GMMs for different spectral periods.

Overall, we propose a short list of three GMMs as models most suitable for induce-seismicity application. Overall, Abrahamson *et al.* (2014), Chiou and Youngs (2014), and Atkinson (2015) GMMs performed better than other models in various frequencies. Atkinson and Assatourians (2017) have also shown that the first two GMMs may be reasonable proxy estimates of median motions from moderate-induced earthquakes in CENA at close distances. The three mentioned models are not specifically developed for CENA, and their generating datasets do not contain data from induced earthquakes. However, these models with appropriate consideration of magnitude and depth scaling in their functional forms outperform CENA-specific models. Another important finding of this study is that the stochastic models favored in the low-seismicity region of CENA appear not to perform better than models developed based on conventional statistical and empirical approaches for induced-seismicity applications.

Data and Resources

The database used for this study was obtained from https://www.sciencebase.gov/catalog/item/57f7d8f2e4b0bc0bec09d04d (last accessed December 2017). Additional figures and tables are provided in the Ⓔ electronic supplement to this article.

Acknowledgments

The authors thank James Kaklamanos and an anonymous reviewer for their constructive and insightful comments that helped us improve this article significantly. This project was partially supported by the Tennessee Department of Transportation.

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.

Al Atik, L. (2015). NGA-East: Ground-motion standard-deviation models for central and eastern North America, *PEER Report Number 2015/07*, Berkeley, California.

Ambraseys, N. N., J. Douglas, S. K. Sarma, and P. M. Smit (2005). Equations for estimation of strong ground motions from shallow crustal earthquakes using data from Europe and the Middle East: Horizontal peak ground acceleration and spectral acceleration, *Bull. Earthq. Eng.* **3,** 1–53.

Ancheta, T. D., R. B. Darragh, J. P. Stewart, E. Seyhan, W. J. Silva, B. S. Chiou, K. E. Wooddell, R. W. Graves, A. R. Kottke, and D. M. Booreet (2014). PEER NGA-West2 Database, *PEER Report Number 2013/03*, University of California, Berkeley.

Atkinson, G. M. (2015). Ground-motion prediction equation for small-to-moderate events at short hypocentral distances, with application to induced-seismicity hazards, *Bull. Seismol. Soc. Am.* **105,** no. 2A, 981–992, doi: 10.1785/0120140142.

Atkinson, G. M., and K. Assatourians (2017). Are ground-motion models derived from natural events applicable to the estimation of expected motions for induced earthquakes? *Seismol. Res. Lett.* **88,** no. 2A, 430–441.

Atkinson, G. M., and M. Morrison (2009). Observations on regional variability in ground-motion amplitudes for small-to-moderate earthquakes in North America, *Bull. Seismol. Soc. Am.* **99,** no. 4, 2393–2409.

Atkinson, G. M., H. Ghofrani, and K. Assatourians (2015). Impact of induced seismicity on the evaluation of seismic hazard: Some preliminary considerations, *Seismol. Res. Lett.* **86,** 1116–1124.

Bindi, D., L. Luzi, F. Pacor, G. Franceschina, and R. R. Castro (2006). Ground-motion predictions from empirical attenuation relationships versus recorded data: The case of the 1997–1998 Umbria-Marches, Central Italy, strong-motion data set, *Bull. Seismol. Soc. Am.* **96,** no. 3, 984–1002.

Bommer, J. J., B. Dost, B. Edwards, P. J. Stafford, J. van Elk, D. Doornhof, and M. Ntinalexis (2016). Developing an application-specific ground-motion model for induced seismicity, *Bull. Seismol. Soc. Am.* **106,** no. 1, 158–173.

Bommer, J. J., P. J. Stafford, J. E. Alarcón, and S. Akkar (2007). The influence of magnitude range on empirical ground-motion prediction, *Bull. Seismol. Soc. Am.* **97,** no. 6, 2152–2170.

Downloaded from https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/4275358/bssa-2017330.1.pdf
by University of Memphis, 16550
on 13 November 2018

Boore, D. M, and K. W. Campbell (2017). Adjusting central and eastern North America ground-motion intensity measures between sites with different reference-rock site conditions, *Bull. Seismol. Soc. Am.* **107,** 132–148.

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. Al 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,** 973–987.

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.

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,** no. 3, 1117–1153.

Douglas, J., B. Edwards, V. Convertito, N. Sharma, A. Tramelli, D. Kraaijpoel, B. M. Cabrera, N. Maercklin, and C. Troise (2013). Predicting ground motion from induced earthquakes in geothermal areas, *Bull. Seismol. Soc. Am.* **103,** no. 3, 1875–1897.

Ellsworth, W. (2013). Injection-induced earthquakes, *Science* **341,** no. 6142, doi: 10.1126/science.1225942.

Farhadi, A., and M. Mousavi (2016). Consideration of the rupture model uncertainties in the probabilistic seismic hazard analysis, *Soil Dynam. Earthq. Eng.* **83,** 191–204.

Frohlich, C., W. Ellsworth, W. Brown, M. Brunt, L. Luetgert, T. MacDonald, and S. Walter (2014). The 17 May 2012 M 4.8 earthquake near Timpson, East Texas: An event possibly triggered by fluid injection, *J. Geophys. Res.* **119,** 581–593.

Goda, K., and G. M. Atkinson (2014). Variation of source-to-site distance for megathrust subduction earthquakes: Effects on ground motion prediction equations, *Earthq. Spectra* **30,** no. 2, 845–866.

Gupta, A., J. W. Baker, and W. Ellsworth (2017). Assessing ground-motion amplitudes and attenuation for small-to-moderate induced and tectonic earthquakes in the central and eastern United States, *Seismol. Res. Lett.* **88,** no. 5, 1379–1389.

Hassani, B., and G. M. Atkinson (2015). Referenced empirical ground-motion model for eastern North America, *Seismol. Res. Lett.* **86,** 477–491.

Hosseini, M., P. Somerville, A. Skarlatoudis, H. K. Thio, and J. Bayless (2016). Constraining shallow subsurface *S* wave velocities with the initial portion of local *P* waves recorded at multiple seismic networks including ANSS and EarthScope Transportable Array in the CEUS, *U.S. Geol. Surv. Final Tech. Rept. Award Number G14AP00110*.

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, doi: 10.1193/070613EQS195M.

Kaklamanos, J., L. G. Baise, and D. M. Boore (2011). Estimating unknown input parameters when implementing the NGA ground-motion prediction equations in engineering practice, *Earthq. Spectra* **27,** 1219–1235.

Keranen, K., H. Savage, G. Abers, and E. Cochran (2013). Potential induced earthquakes in Oklahoma, USA: Links between wastewater injection and the 2011 M_w 5.7 earthquake sequence, *Geology* **41,** 699–702.

Mak, S. (2017). Measuring the performance of ground-motion models: The importance of being independent, *Seismol. Res. Lett.* **85,** no. 5, 1212–1217.

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, doi: 10.1785/0120160232.

Parker, G. A., J. A. Harmon, J. P. Stewart, Y. M. Hashash, A. R. Kottke, E. M. Rathje, W. J. Silva, and K. W. Campbell (2017). Proxy-based V_{S30} estimation in central and eastern North America, *Bull. Seismol. Soc. Am.* **107,** 117–131.

Pezeshk, S., A. Zandieh, K. Campbell, and B. Tavakoli (2015). Ground-motion prediction equations for CENA using the hybrid empirical method with NGA-West2 empirical ground-motion models, in *NGA-East: Median Ground-Motion Models for the Central and Eastern North America Region*, PEER Report Number 2015/04, Chapter 5, Berkeley, California.

Pezeshk, S., A. Zandieh, and B. Tavakoli (2011). Hybrid empirical ground motion prediction equations for eastern North America using NGA models and updated seismological parameters, *Bull. Seismol. Soc. Am.* **101,** no. 4, 1859–1870.

Rennolet, S. B., M. P. Moschetti, E. M. Thompson, and W. L. Yecka (2017). A flatfile of ground motion intensity measurements from induced earthquakes in Oklahoma and Kansas, *Earthq. Spectra* **34,** no. 1, 1–20.

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.

Shahjouei, A., and S. Pezeshk (2016). Alternative hybrid empirical ground-motion model for central and eastern North America using hybrid simulations and NGA-West2 models, *Bull. Seismol. Soc. Am.* **106,** no. 2, 734–754.

Tavakoli, B., F. Sedaghati, and S. Pezeshk (2018). An analytical effective point-source-based distance-conversion approach to mimic the effects of extended faults on seismic hazard assessment, *Bull. Seismol. Soc. Am.* **108,** no. 2, 742–760.

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.

Yenier, E., and G. M. Atkinson (2015). Regionally adjustable generic ground-motion prediction equation based on equivalent point-source simulations: Application to central and eastern North America, *Bull. Seismol. Soc. Am.* **105,** 1989–2009.

Zafarani, H., and A. Farhadi (2017). Testing ground-motion prediction equations against small-to-moderate magnitude data in Iran, *Bull. Seismol. Soc. Am.* **107,** 912–933, doi:10.1785/0120160046.

Zalachoris, G., E. M. Rathje, and J. G. Paine (2017). V_{S30} characterization of Texas, Oklahoma, and Kansas using the *P*-wave seismogram method, *Earthq. Spectra* **33,** no. 3, 943–961.

Department of Civil Engineering
The University of Memphis
Memphis, Tennessee
afarhadi@memphis.edu
spezeshk@memphis.edu
(A.F., S.P.)

Center for Earthquake and Research Information
The University of Memphis
Memphis, Tennessee
nkhshnvs@memphis.edu
(N.K.)

Manuscript received 2 November 2017;
Published Online 3 July 2018

Downloaded from https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/4275358/bssa-2017330.1.pdf
by University of Memphis, 16550
on 13 November 2018