Earthquake Geotechnical Engineering for Protection and Development of Environment and Constructions – Silvestri & Moraci (Eds)
© 2019 Associazione Geotecnica Italiana, Rome, Italy, ISBN 978-0-367-14328-2

Combining ground motion prediction models for epistemic uncertainty minimization

D.Y. Kwak
Hanyang University, Ansan, Gyeonggi-do, South Korea

E. Seyhan
Risk Management Solutions, Inc., Newark, NJ, USA

T. Kishida
Khalifa University, Abu Dhabi, United Arab Emirates

ABSTRACT: Each ground motion prediction equation (GMPE) provides different median ground motion measures and variances computed from a set of input parameters since the data set and methodology used to develop the GMPE vary. These differences are captured by the epistemic uncertainty that can be reduced by combining multiple models. We describe how to minimize the epistemic uncertainty by sensitivity testing on various combinations of four NGA-West2 GMPEs. The correlation levels among models are suggested based on the ranges of moment magnitude, site-to-source distance, site conditions, and selected sub-regions. The prediction errors are highly correlated at short periods among all models, whereas correlations are coarse at long periods. The optimized weight method which uses correlations between errors of models is the most effective to reduce the error variation comparing to other weighting methods. The use of optimized weight method using conditional weights, however, does not significantly further reduce the variation.

1 INTRODUCTION

Ground motion prediction equations (GMPEs) are often used to predict the expected, median ground-motion shaking levels as a function of various model parameters representing the seismic source characteristics, site-to-source distances and site conditions with their associated uncertainties. For a given set of source and site conditions, each GMPE prediction varies because the data set and methodology used to develop the GMPE are different. The ground-motion uncertainty can be attributed to two categories: 1) *aleatory uncertainty* is the inherent variability in the physical system; it is stochastic and cannot be reduced by improving the existing approach; 2) *epistemic uncertainty* is associated with the lack of data and knowledge; it is subjective and can be improved with additional information. The question that is quite commonly asked among researchers and practitioners is how to combine the GMPEs (i.e. GMPE selection and weight assignment) to reduce the epistemic uncertainty to improve our current understanding of the seismic hazard assessment applications. Combining individual models to perform seismic hazard calculations requires weights or ranking of the GMPEs based on the performance of one model over another.

We often combine multiple models linearly. A particular weight, of which the sum is unity, is assigned to each model, and we calculate the weighted sum as the final model prediction. This method is called "linear combination method." Generally, there are three linear combination methods based on weighting schemes:

1. Equal weight (EQ): Assign equally distributed weights to each model. EQ method is often used when models are ranked closely, and only median predictions are available coupled with inadequate knowledge.
2. Inverse-variance weight (INV): The weight is proportional to the inverse of the variance. This method is used when error distributions for each model are available.
3. Optimized weight (OPT): Optimized weights can be determined by minimizing the prediction variance, for which both distributions and correlations among the multiple models are needed.

Above three methods are used in various applications. Seyhan et al. (2014) provide inverse-variance weights to proxy-based models predicting V_{S30} for each region (California, Japan, Taiwan) applied to Next Generation Attenuation (NGA-West2) site database. Kwak et al. (2018) compare how much each linear combination method could reduce the epistemic uncertainty of the combined V_{S30}. Kishida et al. (2018) present optimized weights to existing local magnitude models to predict the moment magnitude using Iranian data sets.

In this study, we applied these linear combination methods to main four NGA-West2 GMPEs to test the possibility of minimization of the epistemic uncertainty. The variance of the residuals and correlation coefficients, which are used for INV and OPT methods, are calculated using the dataset used for the NGA-West2 GMPE development. The model and the dataset are further described in the following section. Correlation coefficients are evaluated from total residuals as well as residuals conditional on key input parameters such as magnitude, source-to-site distance, site conditions, and selected sub-regions to seek the possibility of the further deduction.

2 NGA WEST2 GMPES AND SELECTED DATA SET

2.1 *NGA-West2 GMPEs*

In 2014, the Pacific Earthquake Engineering Research Center (PEER) published a flatfile containing strong-ground motion recordings from shallow crustal earthquakes in active tectonic regions worldwide and five GMPEs as an outcome of a multi-year research program called NGA-West2. Among the five, four main GMPEs that are applicable to broad ranges of site conditions can be listed as Abrahamson et al. (2014), Boore et al. (2014), Campbell and Bozorgnia (2014), and Chiou and Youngs (2014) (hereafter, ASK14, BSSA14, CB14, CY14, respectively). Although each developer team started from the same NGA-West2 flatfile, their data filtering criteria and data selection rules created a variation in the final dataset used for model development. Also, each team uses different ranges of input parameters for the ground motion predictions. The variation in their input parameters selected are shown in Table 1.

The general functional form of median NGA-West2 GMPEs consists of source (f_E), path (f_P), and site (f_S) terms:

$$ln\hat{Y} = f_E + f_P + f_S \tag{1}$$

$$lnY = ln\hat{Y} + R \tag{2}$$

where Y is recorded intensity measure, $\hat{Y}$ is the median prediction, and R is the total residual. While R can be further divided into between-event and within-event residuals, we use R for the calculation of variance and correlation coefficient in this study to compare model-to-model, not event-to-event or site-to-site, relationships.

2.2 *Selected NGA-West2 dataset*

Each of NGA-West2 GMPEs selected ground-motion records filtered by their own criteria for the model development. In this study, we are not concerned with replicating the number of

Table 1. The variation of input parameters adopted for the model development in NGA West2 project.

Parameter	ASK14	BSSA14	CB14	CY14
Magnitude	M_w	M_w	M_w	M_w
Top of Rupture (km)	Z_{tor}	-	Z_{tor}	Z_{tor}
Style of faulting	F_{RV},F_{NM},SS	U, F_{RV},F_{NM},SS	F_{RV},F_{NM},SS	F_{RV},F_{NM},SS
Dip (deg)	Dip	-	Dip	Dip
Down-dip rupture width (km)	W	-	W	-
Closest distance to rupture(km)	R_{rup}	-	R_{rup}	R_{rup}
Hor. dist. to surface proj. (km)	R_{JB}	R_{JB}	R_{JB}	R_{JB}
Hor. dist. from edge of rupture (km)	R_x	-	R_x	R_x
Hor. dist. off end of rupture (km)	R_{y0}	-	-	-
Hanging wall model	F_{HW}	(R_{JB})	F_{HW}	F_{HW}
V_{S30} (m/s)	V_{S30}	V_{S30}	V_{S30}	V_{S30}
V_{S30} for reference rock (m/s)	1100	760	1100	1130
Depth to Vs (km)	$Z_{1.0}$	$Z_{1.0}$ ($dz_{1.0}$)	$Z_{2.5}$	$Z_{1.0}$
Hypocentral depth (km)	-	-	Z_{hyp}	-
Directivity term	-	-	(Z_{hyp})	DDPP
Regional variations	Region	Region	Region	Region
Aftershock factor	F_{AS}	-	-	-

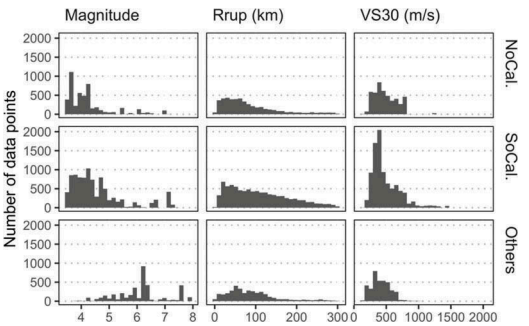

Figure 1. Distribution of selected records conditional on magnitude (**M**), source-to-site distance (R_{rup}), and site condition (V_{S30}) per region. The column represents **M**, R_{rup}, and V_{S30} from 1st to 3rd, and the row represents northern California, southern California, and other regions from 1st to 3rd.

recordings selected by various filtering criteria per each GMPE, thus, we included all the applicable recordings from the flatfile to test if all four models can produce the competitive median predictions using the same input parameters. Figure 1 shows the data histograms of selected records conditional on magnitude (**M**), source-to-site distance (R_{rup}), and time-averaged shear-wave velocity down to 30 m depth of the soil column (V_{S30}) per each sub-region. Parent regions are divided into three sub-regions: Northern California, Southern California, and others. Other regions include many countries worldwide, but some countries only have tens of data points so that we combined them as one category. Distribution shapes of **M**, R_{rup}, and V_{S30} generally follow the log-normal distribution, and it is not region-dependent, except **M** for "others" region. For "others" region, greater **M** events are included in the flatfile so that there are not many records with low **M** events. Note that we apply the lowest usable frequency filter; response spectra with spectral periods higher than the inverse of lowest usable frequency are filtered, so the number of records decreases with periods.

2.3 *Residuals*

Figures 2 a-d show total residuals (R in Eq. 2), mean (μ_R), and standard deviation (σ_R) of R for each GMPE. Since the selected dataset is not the same with the dataset used for the model development, the non-zero μ_R at each spectral period is not surprising. The μ_R is close to zero

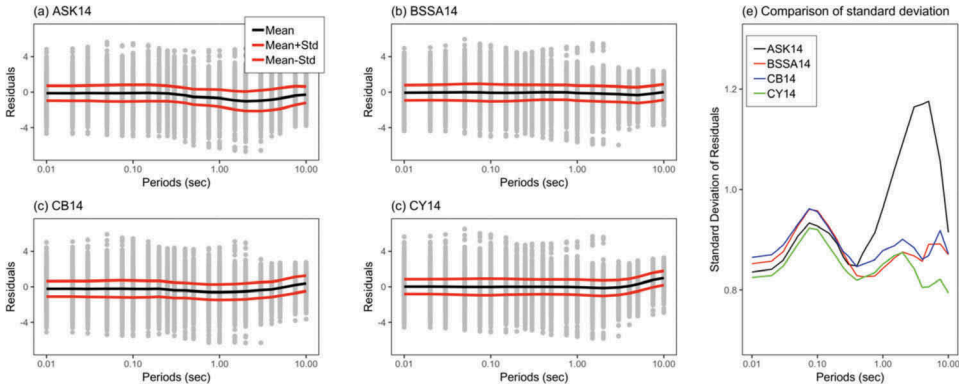

Figure 2. Mean and +/- one standard deviations from mean of residuals. (a) ASK14 (Abrahamson et al. 2014) (b) BSSK14 (Boore et al. 2014) (c) CB14 (Campbell and Bozorgnia 2014) (d) CY14 (Chiou and Youngs 2014). Comparison of standard deviations of four models is shown at (e).

at short periods (< 0.2 sec) consistently for all models, while it becomes larger at long periods. Figure 2e shows the comparison of σ_R of all GMPEs, which are comparable at periods less than 0.2 sec, but the variation of σ_R becomes large at long periods. Herein the GMPE with the lowest σ_R does not necessarily mean that the GMPE is superior over other GMPEs because the selected dataset in this study is different with the one used for the model development, and the dataset used for the comparison is also a sample dataset which does not represent the whole population.

3 CORRELATION OF RESIDUALS

3.1 *Universal correlation coefficients of residuals*

For the optimized weight (OPT) methods, it is necessary to define correlation coefficients (ρ_R) among models. If no dataset were available, it would be difficult to evaluate the correlation coefficients. Using the residuals calculated from NGA-West2 flatfile, ρ_R are evaluated between two GMPEs as shown in Figure 3. The ρ_R are high regardless of model combination for periods less than 0.4 sec. For longer periods (> 0.4 sec), ASK14-CY14 results in the least ρ_R at 2 sec, which is approximately 0.72. The combination of ASK14-BSSA14 and ASK14-CB14 also have relatively lower ρ_R at long periods comparing to the combinations among BSSA14, CB14, and CY14. This indicates that ASK14 is the most independent model from others. The reduction correlation in Figure 3 for ASK14 also corresponds to the increase in σ_R in Figure 2(e). However, this is only in terms of the relativity, and the overall correlation is high among the models.

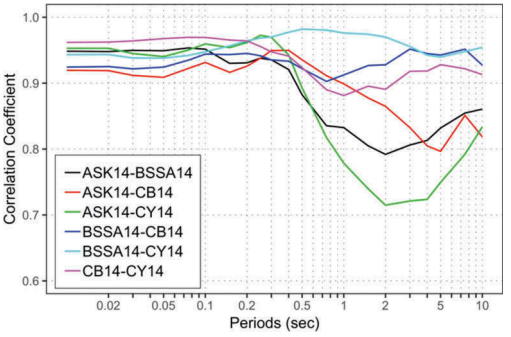

Figure 3. Correlation coefficients between error terms of two NGA West2 GMPEs per period.

3.2 *Conditional correlation coefficients of residuals*

In this section, we evaluate ρ_R between models, which varying by selected input parameters such as **M**, R_{rup}, V_{S30}, and regions to see the possibility of further reduction by selecting optimized weights for each condition.

Figure 4 shows ρ_R from two selected GMPEs for varying **M**, R_{rup}, V_{S30}, and regions, respectively. The **M**, R_{rup}, and V_{S30} are divided as five categories from low to high (**M** < 4, **M** = 4 – 5, **M** = 5 – 6, **M** = 6 – 7, and **M** > 7), close to far (R_{rup} < 30 km, R_{rup} = 30 – 80 km, R_{rup} = 80 – 150 km, R_{rup} = 150 – 250 km, and R_{rup} > 250 km), and slow to fast (V_{S30} < 270 m/s, V_{S30} = 270 – 360 m/s, V_{S30} = 360 – 560 m/s, V_{S30} = 560 – 760 m/s, and V_{S30} > 760 m/s), respectively. Table 2 describes the trends of ρ_R variations per each condition and GMPE combination.

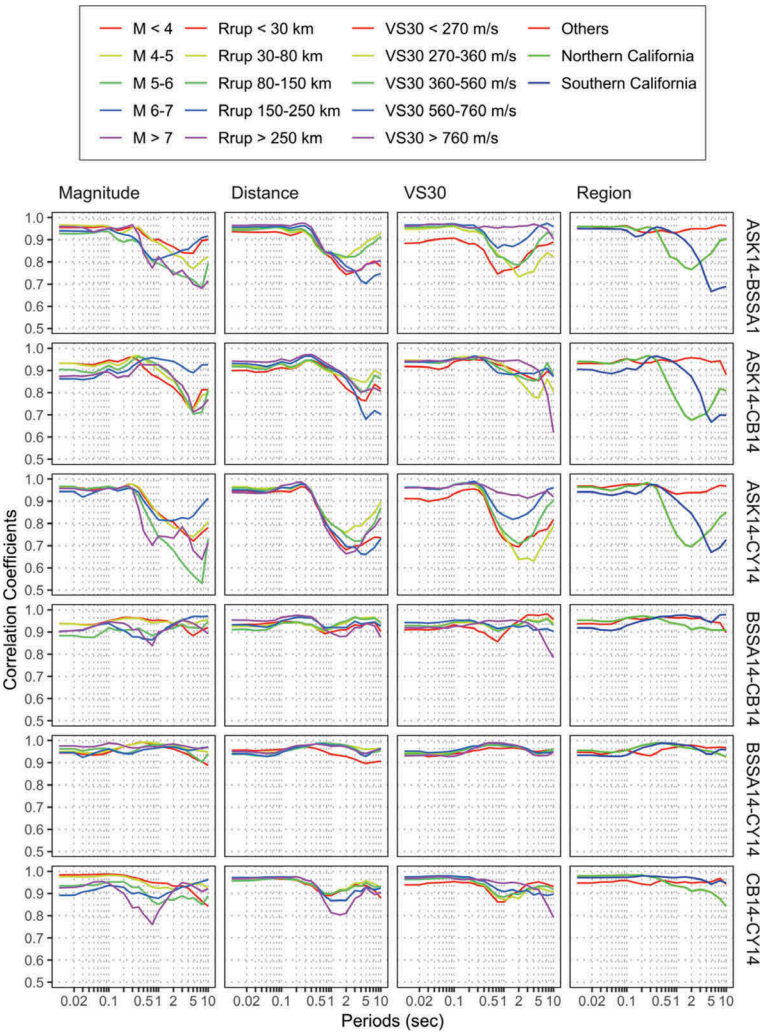

Figure 4. Correlation coefficients between residuals of two selected NGA West2 GMPEs per period. The column represents conditions of magnitude (**M**), source-to-site distance (R_{rup}), V_{S30}, and sub-region from 1st to 4th, and the row represents combination of ASK14-BSSA14, ASK14-CB14, ASK14-CY14, BSSA14-CB14, and CB14-CY14 GMPEs from 1st to 6th. The range of condition is shown in the legend at the top of figures.

Table 2. Descriptions of correlation coefficients variations by magnitude (**M**), source-to-site distance (R_{rup}), site condition (V_{S30}), and selected sub-region per each combination.

Models	Condition	Description
ASK14-BSSA14	**M**	Correlation levels does not vary significantly at short periods (< 0.4 sec). At long periods, small **M** (< 5) tends to have higher correlation levels than large **M** (> 5). The **M** 6-7 has the highest correlation at very long periods.
	R_{rup}	The correlation level variation is not significant by R_{rup} for peridos < 1 sec. The mid-R_{rup} range (30 – 150 km) have higher correlation levels than others at periods > 1 sec.
	V_{S30}	The slow V_{S30} (< 270 m/s) has lower correlation levels at short-to-mid periods. The faster V_{S30} tends to have high correlation at periods > 0.3 sec.
	Region	Northern and southern CA regions have lower correlation levels at periods > 0.2 sec, but the other region has high correlation levels for all periods.
ASK14-CB14	**M**	The **M** 6-7 has the highest correlation among others at periods > 0.6 sec. Others have similar trends.
	R_{rup}	The correlation level variation is not significant by R_{rup}. The mid-R_{rup} range (30 – 150 km) have higher correlation levels than other R_{rup} range at long period (> 2 sec).
	V_{S30}	The correlation of the fast V_{S30} (> 760 m/s) falls down above the periods of 5 sec.
	Region	Northern and southern CA regions have lower correlation levels after 0.2 sec, but other regions have high correlation levels for all periods.
ASK14-CY14	**M**	Correlations start varying above periods > 0.3 sec. The **M** 5-6 has the lowest correlations, and **M** 6-7 has the highest correlations.
	R_{rup}	The correlation level variation is not significant by R_{rup}. The mid-R_{rup} range (30 – 150 km) has higher correlation levels than other R_{rup} range at long period (> 2 sec).
	V_{S30}	The faster V_{S30} tends to have lower correlations for periods > 0.3 sec. The V_{S30} > 760 m/s has lower correlations at short periods as well.
	Region	Northern and southern CA regions have lower correlation levels after 0.2 sec, but other regions have high correlation levels for all periods.
BSSA14-CB14	**M**	Correlation levels are high regardless of **M**.
	R_{rup}	Correlation levels are high regardless of R_{rup}.
	V_{S30}	Correlation levels are high but it becomes lower after 2 sec for fast V_{S30} (> 760 m/s).
	Region	Correlation levels are high regardless of region.
BSSA14-CY14	**M**	Correlation levels are high regardless of **M**.
	R_{rup}	Correlation levels are high regardless of R_{rup}.
	V_{S30}	Correlation levels are high regardless of V_{S30}.
	Region	Correlation levels are high regardless of region.
CB14-CY14	**M**	The high **M** (> 7) has lower correlations at periods of 0.1 – 3 sec.
	R_{rup}	Correlation levels decreases when R_{rup} decreases at mid-period range (0.5-5 sec). Other periods have high correlation levels regardless of R_{rup}.
	V_{S30}	At med-period range (0.4 – 3 sec), the faster V_{S30} tends to have the higher correlations. The correlation for V_{S30} > 760 m/s falls down after 5 sec.
	Region	Correlation levels are high, but it becomes lower after 0.2 sec for northern CA region.

4 REDUCTION OF EPISTEMIC UNCERTAINTY

The standard deviation of the combined GMPE (σ_c) that is the linear summation of n models with weights can be calculated as:

$$\sigma_c^2 = [w_1 \dots w_n] \begin{bmatrix} \sigma_1^2 & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_n^2 \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} \tag{3}$$

where w_i and σ_i^2 are the weight and variance for the i^{th} model, and σ_{ij} is the covariance of i^{th} and j^{th} models. Kwak et al. (2018) describe the estimation of σ_c using three linear combination methods: equivalent weights (EQ), inverse-variance weights (INV), and optimized weights (OPT), which essentially define the weight vector with different scheme as described in the

section *Introduction*. Using variances and correlation coefficients from all residuals of the selected dataset, we find weight vector for each method and evaluate the σ_c as shown in Figure 5. The σ_R from individual GMPEs are also shown. Comparing σ_c among EQ, INV, and OPT methods, the OPT method provides the least σ_c, but the difference from other two methods is minor for short period range (< 1 sec), whereas it becomes greater at long period range (> 1 sec). The maximum reduction of σ_c by OPT method from EQ method is about 8.7% at 5 sec. This is because the correlation levels become relatively low at long periods, for which the OPT method is more efficient. Comparing OPT method with individual GMPE, the CY14 generally provides the least σ_R and the OPT method is slightly lower than CY14. The OPT method provides 1.5% lower σ_c at 2 sec than CY14.

The weights by OPT method in Figure 5 are calculated using all the residuals. We further investigate how much the σ_c could be reduced by assigning different weights to each GMPE based on correlation levels conditional on the model input parameters: **M**, R_{rup}, V_{S30}, and sub-region (Figure 4). The σ_c for this case cannot be simply calculated using Eq. (1) because the weight vector and covariance matrix are varying by each condition. For the σ_c calculation in this case, we followed the following approach:

1. Subtract μ_R (Figure 2) from residuals (R);
2. Find correlation coefficients (ρ) between two particular GMPEs for a subset of R from (1) based on a particular condition;
3. Using ρ_R from (2) and σ_R of all R, find weight vectors using the OPT method;
4. Apply weight vectors from (3), which are dependent on each condition, to GMPE predictions and linearly combine;
5. Calculate R of combined GMPE and estimate σ_c.

In this step, we forced the μ_R to be zero at Step 1, and used total σ_R instead of σ_R estimated from the subset of data at Step 3 assuming that the GMPE prediction is not biased, and the R

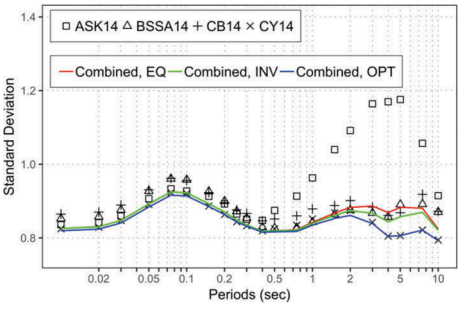

Figure 5. Standard deviations of residuals from individual GMPE and three linear combination methods (EQ: Equal weights; INV: Inverse-variance weights; OPT: Optimized weights). Optimized weights are estimated from correlation coefficients of all residuals between GMPEs.

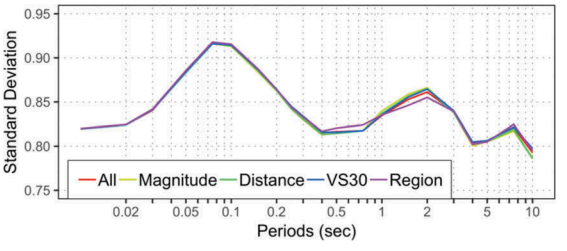

Figure 6. Standard deviations of residuals from optimized weight method using weights defined from correlation coefficients conditional on magnitude, site-to-source distance, V_{S30}, and region.

variation is independent from input parameter. We acknowledge that the *R* variation depends on the GMPE input (Abrahamson et al. 2014, Boore et al. 2014, Campbell and Bozorgnia 2014, Chiou and Youngs 2014), but we did not consider here because we want to see the effect of correlation on the reduction of variation independently.

Figure 6 shows σ_c following above procedure conditioned on **M**, R_{rup}, V_{S30}, and region. The σ_c without any condition (All) is also shown as a reference. As shown, the σ_c from conditioned weights is not decreased from the reference. The maximum improvement is only 0.7 % at 2 sec, which is the case of using ρ conditional on regions. This indicates that the combined GMPE using OPT method without any conditions could be the best-fitted model for linear combination method.

5 CONCLUSIONS

In this study, we investigated the potential use of the linear combination methods to reduce the epistemic uncertainty from four main Next Generation Attenuation (NGA-West2) GMPEs. Among the methods of equal weight (EQ), inverse-variance weight (INV), and optimized weight (OPT), we found that the OPT method provides the least standard deviation (σ_c). However, it is only effective at spectral periods longer than 2 sec. Although the correlations among NGA-West2 GMPEs are generally high (correlation coefficient, $\rho > 0.5$), ASK14 is found to be the most independent from the others relatively. The ρ values vary by the input parameters of the GMPE such as magnitude, source-to-site distance, V_{S30}, and subregion, especially for the case of combination of ASK14 with other GMPEs. The ρ variation is minor among BSSA14, CB14, and CY14. The σ_c reduction was further investigated by combining the GMPEs with different weights conditional on the input parameters, but the reduction was negligible comparing to the result using all the residuals. The combined GMPE with EQ method provides generally less σ_c from the individual GMPE. The 'best' GMPE, at least for the dataset used in this study, gives lower standard deviation, but it is recommended to combine models if the variation information is unknown.

REFERENCES

Abrahamson, N.A., Silva, W.J., & Kamai, R. 2014. Summary of the ASK14 ground motion relation for active crustal regions. *Earthq. Spectra* 30: 1025–1055.

Boore, D.M., Stewart, J.P., Seyhan, E., & Atkinson, G.M. 2014. NGA-West2 equations for predicting PGA, PGV, and 5% damped PSA for shallow crustal earthquakes. *Earthq. Spectra* 30: 1057–1085.

Campbell, K.W. & Bozorgnia, Y. 2014. NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra. *Earthq. Spectra* 30: 1087–1115.

Chiou, B.S.-J. & Youngs, R.R. 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.

Kwak, D.Y., Seyhan, E., & Kishida, T.A. 2018. Method of linear combination of multiple models for epistemic uncertainty minimization. *Proc. 11th Nat. Conf. Earthq. Eng., Earthq. Eng. Res. Inst.*, *Los Angeles*, 25-29 *June 2018*. California: USA.

Kishida, T., Derakhshan, S., Muin, S., Darragh, R.B., Bozorgnia, Y., Kuehn, N., & Kwak, D.Y. 2018. Multivariate conversion of moment magnitude for small-to-moderate magnitude earthquakes in Iran. Earthq. Spectra 34: 313–326.