

Spatial variability of site effects and its correlation with site response in Japan

Cristina Lorenzo-Velazquezⁱ⁾ and Ashly Cabasⁱⁱ⁾

i) Ph.D. Student, Department of Civil, Construction, and Environmental Engineering, North Carolina State University, 915 Partners Way, Raleigh, North Carolina, 27695, United States of America.

ii) Assistant Professor, Department of Civil, Construction, and Environmental Engineering, North Carolina State University, 915 Partners Way, Raleigh, North Carolina, 27695, United States of America.

ABSTRACT

Local soil conditions play an important role in regional seismic hazard assessments due to their influence on earthquake-induced ground shaking and deformation. The different levels of damage and site response at nearby locations correlate to site and geologic conditions variability, as has been reported after past earthquakes. Evaluating spatially variable ground motions (GMs) is key for earthquake reconnaissance efforts and regional seismic hazard assessments. This study focuses on the evaluation of spatial correlations in site parameters (e.g. time-averaged shear-wave velocity to a depth of 30 meters) at Kiban-Kyoshin Network (KiK-net), and their comparison to the observed spatial correlation of the residuals from ground motion intensity measures (IMs) from the Mw 9.1 Tohoku earthquake. Current spatial correlation models treat site effects either as a fixed amplification factor or as randomized amplifications, but site effects are neither fixed nor random. Hence, geostatistical methods are used here to estimate spatial correlations between parameters that control site response and integrate their effects on resulting spatially variable ground motions. In this work, we evaluate the significance of the spatial correlation for different site parameters with respect to the GM amplification IMs residuals.

Keywords: site response, ground motions, spectral acceleration, spatial correlations

1 INTRODUCTION

Site response to earthquake ground motions varies spatially leading to a range of damage in regions, such as deformations. Previous studies hypothesize that soil condition heterogeneity may influence spatial correlation in a region (Jayaram and Baker, 2009; Sokolov et al., 2012) and that both site effects and path effects may cause event-to-event variability (Sokolov et al., 2012; Bodenmann et al. 2023). Therefore, the spatial correlation of site data can inform how ground motion (GM) intensity measures (IMs) correlates spatially.

Our case study location is Japan as it has significant spatial coverage of strong ground motion recording stations and regionally available site parameters. Specifically, the March 11, 2011 Mw 9.1 Tohoku earthquake, is used to investigate how site parameters relate to the site response regionally due to a high-magnitude earthquake. First, the spatial correlation for the commonly used site parameter, time-average shear wave velocity to a depth of 30 meters (V_{S30}) was calculated. Then, we calculated spatial correlations based on within-events residuals derived from predicted GM IMs using a ground motion model (GMM) for subduction zones. These to improve our understanding

of local site conditions and regional site amplification patterns.

2 DATA

2.1 Event Data

To study the spatial distribution of site parameters and site response, we have selected 239 stations with a rupture distance (R_{rup}) less than 300 meters and both, observed spectral accelerations (SA) and measured V_{S30} values available (Fig. 1). We used Bahrapouri and Rodriguez-Marek (2021) database, which contains computed intensity measures such as SA at multiple periods, average shear-wave velocity to a depth of 30 meters (V_{S30}), and source to site distances (R_{rup}), among other GM parameters for the Kiban-Kyoshin network (KiK-net) stations. Spectral accelerations at 0.01 and 1 seconds were chosen to examine the correlation between site parameters and SA at short and long periods.

2.2 Site Conditions

Large regional geologic structures such as volcanic arcs can exert an influence on ground motions as attenuation properties of the rock formations near the volcanic arc can vary (e.g., Pei et al. 2009). In Japan, the influence of this geologic structure has to be considered

when studying ground motions recorded on stations located to the east (forearc stations) and west (backarc stations) of the volcanic front (e.g., Cabas et al. 2017). From the 239 KiK-net stations selected, 148 are backarc stations and 91 are forearc stations.

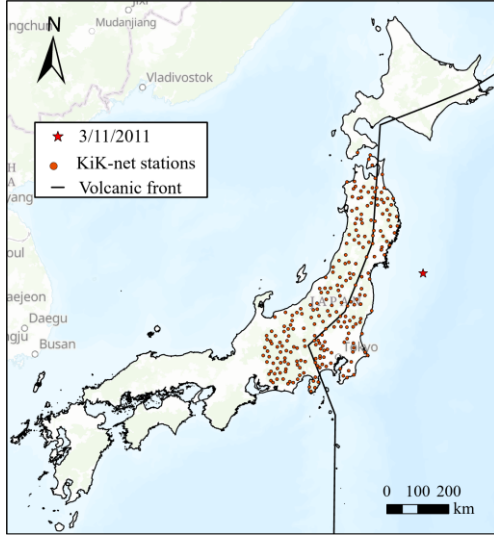


Fig. 1. Location of the Tohoku Mw 9.1 earthquake epicenter, the 239 KiK-net stations selected for this study, and the Japanese volcanic front.

In addition, V_{S30} is considered as a site parameter in this work because it is widely used to capture site effects within GMMs. We examine potential differences in the resulting spatial correlation structure when using measured values of V_s to compute V_{S30} and proxy-based V_{S30} estimates as a function of topographic slope (e.g., Wald and Allen 2007). Considering the limited availability of measured V_s profiles at strong ground motion stations worldwide, proxy-based methods have been used to complement incomplete site metadata in the development of GMMs and in other applications. In this work, we hypothesize that spatial correlations of ground motion intensity measures obtained via GMMs that parametrize site effects using V_{S30} can be limiting for two reasons: 1) V_{S30} only characterizes the very near-surface material stiffness and does not provide information about deep geologic structures or attenuation properties, and 2) mixing measured and proxy-based V_s values may not fully represent the real spatial correlation structure among sedimentary deposits that share common geology and/or depositional environments. Testing the aforementioned hypotheses could provide a path forward to reduce uncertainties in current correlation models and GMMs predictions. In this work, we use the Wald and Allen (2007) topographic-slope proxy-based V_{S30} values for Japan. These proxy-based V_{S30} values result from a global model built using topographic slope gradient at 30 arcsec resolution and collected V_{S30} values from active tectonic regions and stable continental regions.

This case study provides an opportunity to investigate how correlation structures vary when

utilizing diverse sources of V_{S30} data available at all recording stations in conjunction with GM IMs from the event of interest. Fig. 2a, presents a histogram with both measured and proxy-based V_{S30} , including the site classes of the National Earthquake Hazards Reduction Program (NEHRP). As shown in Fig. 2a, the bimodal distribution of the proxy-based V_{S30} values is not representative of the distribution of real measurements, which often follows a lognormal distribution. The majority of our study sites are within NEHRP site classes C (very dense soil or soft rock) and D (stiff soil). KiK-net stations are known to be located at stiffer sites, which could lead to an inherent bias when studying the site response relative to these sites.

Fig. 2b, compares the measured V_{S30} values to the topographic slope proxy-based V_{S30} (Wald and Allen, 2007) values. The scatter presented in Fig. 2b, shows how the use of the proxy-based V_{S30} can add a systematic bias into GMMs. For example, Fig. 2b illustrates how measured V_{S30} values ranging from 307 to 1433 m/s may correspond to a single estimated value of proxy-based V_{S30} equal to 775 m/s, which could lead to an under- or overestimation of the GM amplification. The majority of the data presented in Fig. 2b presents differences larger than 20%, only the 35% (84) stations fall within a $\pm 20\%$ difference (see dashed lines in Fig. 2b).

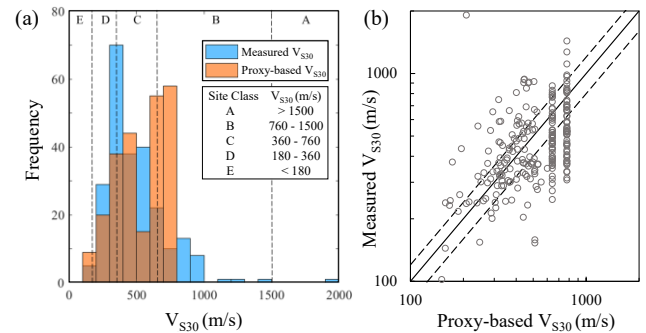


Fig. 2. (a) Histogram of measured and proxy-based V_{S30} values, including the National Earthquake Hazards Reduction Program (NEHRP) site classes. (b) Comparison of measured V_{S30} values and topographic slope proxy-based V_{S30} with a 1:1 slope (solid line) and a threshold of a $\pm 20\%$ offset (dashed lines) are included for reference.

3 METHODOLOGY

We study the spatial correlation of site terms (i.e., V_{S30} measured and from the proxy-based model) and GM IMs using geostatistics. Particularly, a residual analysis was performed using a GMM to examine the GM IMs correlation at different periods.

3.1 Ground motion model

The 2011 Mw 9.1 Tohoku earthquake occurred within the northeast Japan subduction zone, therefore the Abrahamson et al. (2016) GMM for horizontal component response spectral values is used.

Ground motions recorded from an earthquake are influenced by source, path, and site effects. All of those

effects are also considered in GMMs with the use of some parameters such as moment magnitudes, rupture distances (R_{rup}), and V_{S30} , among other terms. The latter is the only parameter that considers the surficial geology that may affect ground motion amplification along with backarc and forearc site identification. In Japan, these subregions (i.e., backarc and forearc) can exhibit a significant difference in amplitude and attenuation characteristics. According to Abrahamson et al. (2016), this GMM captures attenuation between backarc and forearc sites as no trend with distance is observed.

GMMs compute a predicted intensity measure such as spectral acceleration (SA) as follows:

$$\ln SA_{ij} = \mu_{lnSA}(rup_i, site_j) + R_{ij} \quad (1)$$

where the natural logarithm of SA_{ij} is the observed spectral acceleration at a specific period due to earthquake rupture i at site j ; $\mu_{lnSA}(rup_i, site_j)$ is the predicted mean $\ln SA$ value, and R_{ij} is the total residual.

3.2 Residual analysis

In this study, residuals relative to the aforementioned GMM are used to remove the effects of attenuation with distance, as strong signals interfere with the stationarity of the data for an adequate correlation estimation. The total residual was calculated (equation 2) with the geometric mean using the SA of the two orthogonal horizontal components from Bahrapouri and Rodriguez-Marek (2021) database and the predicted SA using Abrahamson et al. (2016).

$$R_{ij} = \ln SA_{ij} - \mu_{lnSA}(rup_i, site_j) \quad (2)$$

Total residuals (R ; equation 3) are partitioned into between-event (inter-event; δB) and within-event (intra-event; δW) residuals (Al Atik et al. 2010).

$$R_{ij} = \delta B_i + \delta W_{ij} \quad (3)$$

The δB residual accounts for the systematic source specific effects, which are the same value across all the stations as they depend on the rupture and not to a specific site. For a substantial number of records for a given earthquake it approaches the arithmetic mean of the total residuals ($\bar{R}$) according to Baker et al. (2021). We used $\bar{R}$ to obtain statistically independent data by calculating the within-event residual:

$$\delta W_{ij} = R_{ij} - \bar{R} \quad (4)$$

The within-event residual contains the systematic site and path effects, varying per site. This residual has been used extensively to estimate spatial correlations in previous studies (Chen et al. 2021; Kuehn and Abrahamson, 2020).

3.4 Spatial Correlations

Geostatistical approaches have been applied to identify spatial patterns in the data and to develop frameworks to work on the damage mitigation after large

earthquakes. To estimate the spatial correlation of the parameters under study, we use the semivariogram. The latter captures how closer observations are more similar to each other than more widely separated observations. This means that the model main assumptions are that correlations are stationary and isotropic (i.e., all sites separated by the same distance have the same semivariance). For the within-event residuals from an event, i at two stations (j and $j+h$) separated by a distance h :

$$\gamma(h) = \frac{1}{2} [\delta W_{ij} - \delta W_{ij+h}]^2 \quad (5)$$

The empirical semivariogram is obtained by estimating the semivariance of the observations at sites with the same distance as:

$$\hat{\gamma}(h) = \frac{1}{2n(h)} \sum_{d(j,j+h)=h} [\delta W_{ij} - \delta W_{ij+h}]^2 \quad (6)$$

The spatial correlation is evaluated by fitting a model to the empirical semivariogram. There are a variety of models that are used to provide a mathematical function to capture the relationship between values and distances. For our study, we chose the exponential model because it was found to provide the best fit to the empirical semivariogram and it has been used for GM residual analyses in previous studies (e.g., Baker and Chen, 2020 and Baker et al., 2021).

$$\gamma(h) = s \left(1 - \exp\left(-\frac{3h}{r}\right) \right) \quad (7)$$

where h is the separation distance, s the sill, and r the range. Beyond the range, where the variation (sill) levels reaches a plateau, the correlation is effectively reduced.

4 RESULTS AND DISCUSSION

Our analyses provide an insight on the potential implications of using a measured or proxy-based site parameter as we calculate spatial correlations of GM IMs. In addition, the dependencies of site parameters and GM IMs from records corresponding to a large magnitude earthquake, such as the Mw 9.1 Tohoku earthquake were examined.

Comparisons of the variation of predicted SA values using measured or proxy-based V_{S30} values (Fig. 3) show that the use of different V_{S30} values in the GMMs would influence predicted SA at longer periods. Predictions for SA at 3 seconds were also tested and a similar scatter to that observed for SA at 1 second was obtained. This particular observation is relevant when studying the spatial correlation based on proxy-based values of IMs at longer periods. The use of only measured, only proxy-based V_{S30} , or mixing V_{S30} values from different sources could affect the correlation of IMs differently for different periods. It must be noted that GMs IMs at long periods are no longer controlled by near surface geologic structures, which could also explain the larger scatter observed in Fig. 3b (beyond the known differences in

V_{S30} observed in Fig. 2).

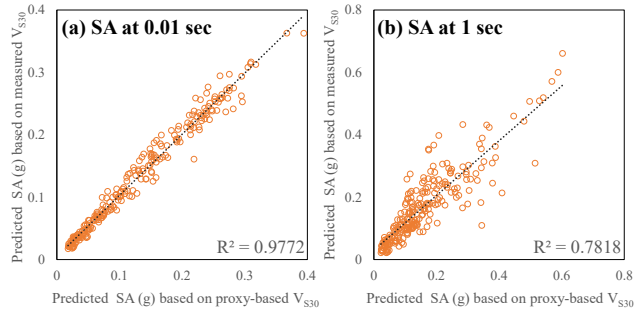


Fig. 3. Dependence of predicted SA at (a) 0.01 second and (b) 1 second using measured and proxy-based V_{S30} .

Fig. 4, presents the distribution of the within-event residuals computed with equation 4 using the observed and predicted SA values at periods of 0.01 and 1 second. Negative values represent overprediction, whereas positive values indicate underprediction of the GM IMs. In addition, the distributions of measured V_{S30} and geologic groups (by geologic age) are presented in Fig. 4c and d, respectively. The geologic groups include Quaternary Holocene (QH), Pleistocene (QP), Tertiary upper (TU) and lower (TL), and Cretaceous (K).

Fig. 4a depicts a systematic underprediction for the

within-event residuals (i.e., positive residuals for the short period SA) at forearc stations closer to the epicenter and with similar paths. This underestimation is more pronounced in the within-event residuals recorded at short-period SA ($T = 0.01$). That region has zones with lower V_{S30} and younger deposits but also pockets showing the presence of outcropping older rock formations (e.g., from the Cretaceous and older deposits). However, more scatter is observed in the spatial distribution of measured V_{S30} values and the geologic groups. On the other hand, the distribution of within-event residuals for SA at 0.01 and 1 seconds (Fig. 4a and 4b, respectively) when crossing the volcanic front are overall depicting an overprediction of the GM IM (i.e., negative residuals in blue). For seismic waves that cross the volcanic belt, reaching backarc sites a stronger attenuation of the ground motion is expected as waves travel through the volcanic front (Ghofrani and Atkison, 2011). Additionally, towards the south region of Japan, where clusters of negative residuals are also apparent, Fig. 4d shows the presence of mostly older geologic formations, yet not consistently high values of V_{S30} (Fig. 4c). More information on the subsurface is needed in these areas to fully understand what is driving the bias observed in the residuals.

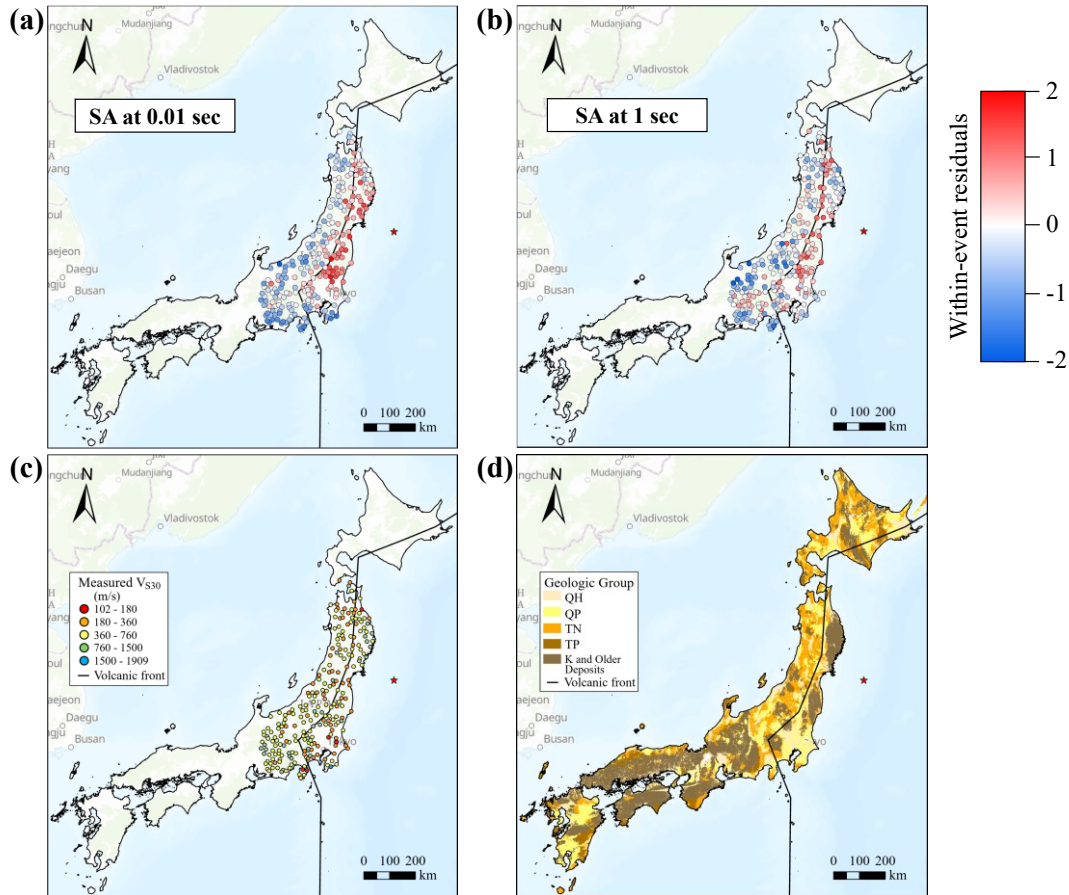


Fig. 4. Distribution for within-event residuals (δW) for SA at (a) 0.01 second and (b) 1 second. The distribution of (c) measured V_{S30} (presented ranges are based on the NEHRP site classes and the minimum and maximum values are constrained to those from the measured V_{S30} values) and (d) geologic groups (where darker colors represent older deposits) obtained from the Geological Survey of Japan (2022).

Fig. 5 presents the within-event residuals for SA at 0.01 and 1 sec against the corresponding V_{S30} measured values. The binned means and error bars (representing $\pm$ one standard deviation) are included to capture any potential trends in the residuals. As seen in Fig. 5, there is a downward-sloping trend that indicates an underprediction for lower V_{S30} values (i.e., less than 450 m/s) and an overprediction with increasing V_{S30} values. This trend is more pronounced for the shorter period SA (Fig. 5a).

Although V_{S30} is used for site classification, sites with the same V_{S30} can have different deeper Vs profiles, leading to a different site amplification (Atik et al. 2010). Additionally, there is a potential temporal variation of V_{S30} (Bonilla et al. 2019) with the onset of nonlinear behavior and the associated larger shear strains that the subsurface materials are possibly subjected to the strong shaking associated with the Tohoku earthquake.

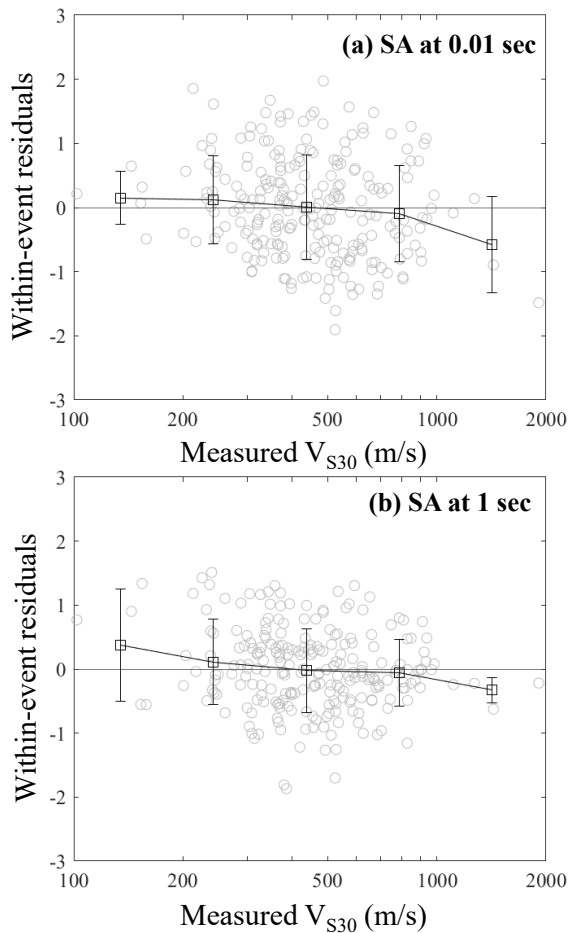


Fig. 5. Dependence of measured V_{S30} for Mw 9.1 Tohoku's within-event residuals for SA at (a) $T = 0.01$ second and (b) $T = 1$ second. For both the binned means and error bars (representing $\pm$ standard deviation).

Typically, correlation distances of SA range within two-digit values (i.e., tens of km), as seen in previous studies (Jayaram and Baker, 2009). However, the obtained correlation lengths (i.e., the range from the semivariogram; r in equation 7) in Table 1 are in the

order of three-digit values (i.e., hundreds of km). One potential reason could be the onset of nonlinearity induced by the intense ground motions generated by the Mw 9.1 earthquake, which could result in SA at sites separated by longer distances still highly correlated. However, more analyses are needed to test this hypothesis.

The correlation lengths (i.e., range) obtained for SA decreased with increasing oscillator period, as seen in Table 1, where the range for SA at 0.01 s is 323 km, while its counterpart for SA at 1 s is 178 km. Previous studies have shown that with increasing SA periods, the range of correlation increases (Jayaram and Baker, 2009). Hence, further investigation is required to fully understand the mechanisms driving this different pattern in our data.

Table 1. Ranges for δW based on the residuals calculated using the predicted SA values from Abrahamson et al. (2016) considering measured and proxy-based V_{S30} . Ranges for measured and proxy-based V_{S30} values are also presented.

Parameter	Range (km)	
	Measured V_{S30}	Proxy-based V_{S30}
δW [SA at 0.01 s]	323	314
δW [SA at 1 s]	178	239
Measured V_{S30}	250	
Proxy-based V_{S30}	287	

One implication of the use of within-event residuals is that they include systematic path and site effects (Atik et al. 2010). Additionally, our observations could be attributed to the specific event under study, as stationarity was assumed by calculating the semivariogram of residuals of this single event. A more robust analysis will follow to move from a stationary to a non-stationary analysis. This will be accomplished by analyzing earthquakes from the same tectonic setting with close-distance epicenters and using a site-specific approach (Chen et al. 2021). Adding more events, specifically lower magnitude foreshocks and aftershocks from the 2011 Tohoku sequence, will also contribute to the investigation of the potential temporal variation of V_{S30} after large events.

5 CONCLUSION

In this study, the spatial correlations of the site and GM IMs are analyzed to further understand if the spatial correlation of the site parameters could inform the spatial correlations of earthquake GM IMs and reduce the variability in GMMs. It is typically assumed that spatial correlations of GM IMs reduce with increasing separation distance, but systematic patterns driven by geology and/or depositional environments can help advance our characterization of spatial correlations beyond the assumptions of stationarity and isotropy. In this study, the correlation structure of a short- and long-period SA as well as measured and proxy-based V_{S30} were modeled using semivariograms.

The implications of using measured or proxy-based V_{S30} for estimating GM IMs using GMMs were assessed. We found that this choice is particularly important when selecting a correlation model for GM IMs at longer periods. Our results in terms of SA at short periods using measured V_{S30} values in the GMM correlated well with their counterparts obtained using proxy-based V_{S30} values.

The distribution of within-event residuals for the SA at 0.01 s and 1 s obtained for the Mw 9.1 Tohoku earthquake shows a systematic overprediction for the sites crossing the volcanic front, as well as clusters of positive residuals (underprediction) in the forearc region. Observed patterns were associated with the potential triggering of nonlinearity and differences in attenuation properties affecting GMs crossing the volcanic belt. Moreover, our correlation lengths for SA presented a reduction at long periods. However, previous studies have shown that when longer periods are studied, the correlation length is expected to increase from that obtained at shorter periods. This is attributed to longer-period SA not being as affected by surficial heterogeneities compared to shorter-period SA.

A more rigorous analysis is required to accurately capture site-specific correlations while isolating the influence of both source and path effects on correlation structures. Our next steps include studying the correlation structures and the dependencies with GM IMs of other site response parameters that can capture deeper geologic structures, such as the site fundamental period (f_p) and depth to V_{SX} horizon (Z_{SX}), as well as the patterns obtained for co-located events of different magnitudes.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-2137100 and the NSF CAREER Grant No. CMMI-2145466. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

- 1) Abrahamson, N., Gregor, N., and Addo, K. (2016): BC Hydro Ground Motion Prediction Equations for Subduction Earthquakes. *Earthquake Spectra*, 32(1), 23-44. <https://doi.org/10.1193/051712EQS188MR>
- 2) Atik, L. A., Abrahamson, N., Bommer, J. J., Scherbaum, F., Cotton, F., and Kuehn, N. (2010): The variability of ground-motion prediction models and its components. *Seismological Research Letters*, 81(5), 794-801.
- 3) Bahrapouri, M., and A. Rodriguez-Marek. (2021): Ground Motion Parameters for KiK-net Records: An Updated Database. *DesignSafe-CI*. <https://doi.org/10.17603/ds2-c0ts-c070>
- 4) Baker, J. W., Bradley, B. A., and Stafford, P. J. (2021): *Seismic Hazard and Risk Analysis*. Cambridge University Press, Cambridge, England.
- 5) Baker, J.W., and Chen, Y. (2020): Ground motion spatial correlation fitting methods and estimation uncertainty. *Earthquake Engineering and Structural Dynamics*, 49, 1662–1681. <https://doi.org/10.1002/eqe.3322>
- 6) Bodenmann, L., Baker, J. W., and Stojadinović, B. (2023): Accounting for path and site effects in spatial ground-motion correlation models using Bayesian inference, *Natural Hazards and Earth System Sciences*, 23(7), 2387-2402.
- 7) Bonilla, L. F., Guéguen, P., and Ben-Zion, Y. (2019): Monitoring coseismic temporal changes of shallow material during strong ground motion with interferometry and autocorrelation. *Bulletin of the Seismological Society of America*, 109(1), 187-198.
- 8) Cabas, A., Rodriguez-Marek, A., & Bonilla, L. F. (2017): Estimation of site-specific kappa (κ_0)-consistent damping values at KiK-net sites to assess the discrepancy between laboratory-based damping models and observed attenuation (of seismic waves) in the field. *Bulletin of the Seismological Society of America*, 107(5), 2258-2271.
- 9) Chen, Y., Bradley, B. A., and Baker, J. W. (2021): Nonstationary spatial correlation in New Zealand strong ground-motion data. *Earthquake Engineering and Structural Dynamics*, 50(13), 3421-3440.
- 10) Geological Survey of Japan (2022): Seamless digital geological map of Japan 1: 200,000. March 11, 2022 version 2. Geological Survey of Japan, National Institute of Advanced Industrial Science and Technology. Available at: <https://gbank.gsj.jp/seamless/use.html> (accessed January 24, 2023).
- 11) Ghofrani, H., and G. M. Atkinson, (2011): Forearc versus Backarc Attenuation of Earthquake Ground Motion. *Bulletin of the Seismological Society of America*, 101(6), 3032-3045, doi: 10.1785/0120110067.
- 12) Jayaram, N., & Baker, J. W. (2009). Correlation model for spatially distributed ground-motion intensities. *Earthquake Engineering & Structural Dynamics*, 1109-1131. DOI: 10.1002/eqe.922
- 13) Kuehn, N. M., and Abrahamson, N. A. (2020): Spatial correlations of ground motion for non-ergodic seismic hazard analysis. *Earthquake Engineering & Structural Dynamics*, 49(1), 4-23. DOI: 10.1002/eqe.3221
- 14) Pei, S., Z. Cui, Y. Sun, M. N. Toksöz, C. A. Rowe, X. Gao, J. Zhao, H. Liu, J. He, and F. D. Morgan (2009): Structure of the upper crust in Japan from S-wave attenuation tomography, *Bulletin of the Seismological Society of America*, 99(1), 428–434.
- 15) Sokolov V, Wenzel F, Wen KL, and Jean WY. (2012): On the influence of site conditions and earthquake magnitude on ground-motion withinearthquake correlation: analysis of PGA data from TSMIP (Taiwan) network. *Bulletin of the Seismological Society of America*, 10(5):1401-1429. DOI: 10.1007/s10518-012-9368-5
- 16) Wald, D.J. and Allen, T.I. (2007): Topographic slope as a proxy for seismic site conditions and amplification. *Bulletin of the Seismological Society of America*, 97, 1379–1395.