Performance evaluation of conditional generative modeling for ground motion using NGA-West2 Database

Rie Nakata[1,2], Michael W. Mahoney[1,2,3], Pu Ren[2], Nori Nakata[1,2,4], Maxime Lacour[1,3]
[1]International Computer Science Institute [2]Lawrence Berkeley National Laboratory [3]Department of Statistics, University of California, Berkeley [4]Massacchusetts Institute of Technology

Introduction

Generative artificial intelligence (GenAI) is emerging as a promising approach for modeling ground motion waveforms and intensities, with potential applications in non-ergodic seismic hazard analysis (Wang et al., 2021; Florez et al., 2022; Shi et al., 2024; Ren et al., 2024). These models learn complex representations of source, path, and site effects directly from data, offering a flexible alternative to traditional empirical ground motion models (GMMs). While recent studies have demonstrated the feasibility of GenAI for ground motion prediction, their evaluation within established seismological frameworks remains limited.

We develop a conditional generative modeling (CGM) framework for ground motion prediction. Previous work under SCEC support (project 24123) demonstrated its capability to reproduce observed ground motion behavior and spatial variability in the San Francisco Bay Area.

In this study, we evaluate the performance of CGM-based models using the NGA-West2 database (Ancheta et al., 2014), a widely used benchmark for ground motion prediction. The analysis focuses on benchmarking CGM predictions against established empirical GMMs in terms of median ground motion, variability, and scaling behavior with magnitude, distance, and site conditions.

Method

We implement the conditional generative modeling (CGM) framework using conditional variational autoencoder (VAE) models, as shown in Figure 1, building on our previous work. VAEs are selected due to their training stability, lightweight architecture, and robust optimization behavior.

The VAE model learns a low-dimensional representation of ground motion data and enables stochastic generation of new samples. It extends the traditional autoencoder (AE), which compresses input data into a deterministic latent representation and reconstructs the input from this compressed space. In contrast, the VAE adopts a probabilistic formulation, where the latent space represents a distribution rather than a single deterministic encoding. A Gaussian distribution, parameterized by mean μ and standard deviation σ, is used to represent this latent space.

We use the NGA-West2 database (Ancheta et al., 2014) to train and evaluate the CGM models. The database contains 21,336 three-component recordings from approximately 600 shallow crustal earthquakes. Established NGA-West2 ergodic GMMs provide predictions of ground motion intensity measures such as PGA, PGV, and PSA, and serve as benchmarks for model evaluation.

In this study, we develop two CGM-based models to emulate representative NGA-West2 GMMs: the ASK14 model of Abrahamson et al. (2014), and the BSSA14 model of Boore et al. (2013), referred to as CGM-ASK and CGM-BSSA. The models are trained using the same input parameters and data subsets as the corresponding empirical GMMs (see Table 1). To maintain consistency with the NGA-West2 framework, the models are designed to directly generate ground motion intensity measures. Model performance is evaluated in terms of median ground motion prediction and aleatory variability, as well as fundamental physical behaviors including scaling with magnitude, rupture distance, and site conditions.

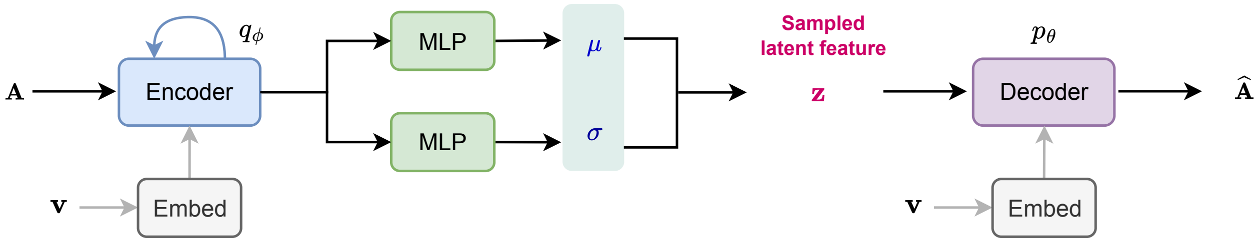

Figure 1: Proposed CGM-ASK and CGM-BSSA architectures. The VAE comprises two key components: an encoder; and a decoder. The encoder first compresses the input data A, and then a nonlinear neural network (a multilayer perceptron, MLP) to extract latent features and output the mean, μ, and standard deviation, σ, of the Gaussian distribution. The decoder then reconstructs the input variable from the latent variable z sampled from the probabilistic latent space, and thus the output from VAE is a probability distribution, instead of a fixed (deterministic) value. Conditional variables are incorporated to the CGM framework through contrastive learning-based embeddings.

	ASK	BSSA
Moment magnitude	M	M
Depth to top of rupture (km)	Z_{TOR}	-
Dip	δ	-
Down-dip rupture width (km)	W	-
Closest distance to rupture plane (km)	R_{RUP}	-
Horizontal distance to surface project of rupture plane (km)	R_{JB}	R_{JB}
Horizontal distance to top edge of rupture (km)	R_X	-
Average shear-wave velocity in top 30 m (m/s)	V_{S30}	V_{S30}
Depth to 1.0 km/s boundary (km)	$Z_{1.0}$	$\delta Z_{1.0}$
Applicable Magnitudes	3-8.5	3-8.5
Applicable distance (km)	0-300	0-400
Applicable V_{S30}	180-1500	150-1500
Applicable period range	PGA-10 s, PGV	PGA-10 s, PGV

Table 1: Summary of input parameters and applicable range of the ASK and BSSA models. Modified from (Gregor et al., 2014).

Results

We evaluate the performance of the CGM-ASK and CGM-BSSA models using the NGA-West2 database following the framework of (Gregor et al., 2014). The analysis focuses on median ground motion prediction, aleatory variability, and scaling behavior with respect to R_{JB} distance, V_{S30}, and magnitude. For each model, 100 samples are generated for a given earthquake and site condition to estimate the median ground motion.

The ability of the models to reproduce spectral shape is first examined. Figure 2 shows the response spectra for periods between 0.01 and 1.0 s for hard rock conditions ($V_{S30} = 800$ m/s). Results are compared between CGM-BSSA, CGM-ASK, and the corresponding empirical models (BSSA14 and ASK14) for $M = 5$, 6.7, and 8, along with the median of the observations. Both CGM models reproduce the overall shape of the response spectra predicted by the empirical GMMs. The agreement is generally stronger for the BSSA-based parameterization than for ASK14. This difference is attributed to the larger number of conditioning variables in the ASK parameterization, which increases model complexity and may introduce non-uniqueness in the learned mapping between inputs and ground motion.

We next examine variability characteristics of the models. Figure 3 compares the aleatory variability and the standard deviation of within-event residuals. The aleatory variability of the CGM models is estimated from the standard deviation of the 100 generated samples. The resulting variability is comparable in magnitude to that of the empirical GMMs, but exhibits a stronger dependence on spectral period. Residuals are defined as $\ln(Y_{\text{obs}}) - \ln(Y_{\text{pred}})$. The standard deviation of the within-event residuals is smaller for the CGM models than for BSSA14 and ASK14, likely reflecting the increased flexibility of the neural network models.

The scaling behavior of ground motion with respect to site conditions, distance, and magnitude is summarized in Figure 4. The CGM models reproduce the observed dependence on V_{S30} across spectral periods of 0.01, 0.2, 1.0, and 3.0 s, showing consistency with the trends predicted by the empirical GMMs.

The models also capture the distance attenuation behavior, including the saturation of ground motion at short distances observed in NGA-West2 GMMs. However, the number of observations decreases significantly for $R_{JB} < 10$ km, which introduces challenges for model generalization in this regime. For rock sites, the CGM predictions tend to be slightly lower than the empirical models, although they generally remain within one standard deviation.

The magnitude scaling behavior shows that the CGM models capture the expected saturation of ground motion with increasing magnitude, particularly at larger distances. The limited number of large-magnitude observations may affect the robustness of model generalization in this regime, with a more pronounced saturation observed in the CGM-BSSA model than in CGM-ASK.

Conclusions

We evaluate the performance of conditional generative models (CGM-ASK and CGM-BSSA) for predicting ground motion intensities using the NGA-West2 database. The CGM models reproduce key characteristics of empirical GMMs, including median response spectra, scaling with distance, magnitude, and site conditions, and overall levels of aleatory variability. Differences between CGM-ASK and CGM-BSSA highlight the impact of conditioning complexity on model behavior, with increased dimensionality potentially introducing non-uniqueness in the learned relationships. The models also exhibit sensitivity to data density, particularly at short distances ($R_{JB} < 10$ km) and for large-magnitude events, where limited observations affect generalization. Overall, the results demonstrate that CGM provides a viable framework for ground motion modeling, while emphasizing the importance of data coverage and model design for robust performance.

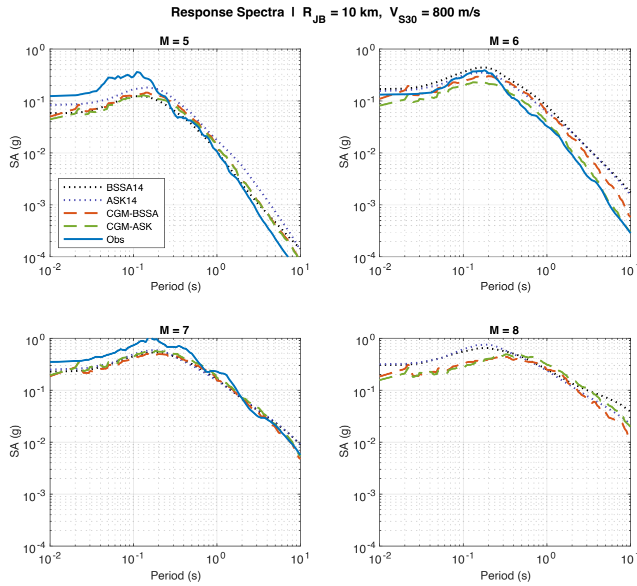

Figure 2: Median response spectra for $M = 5$–8.

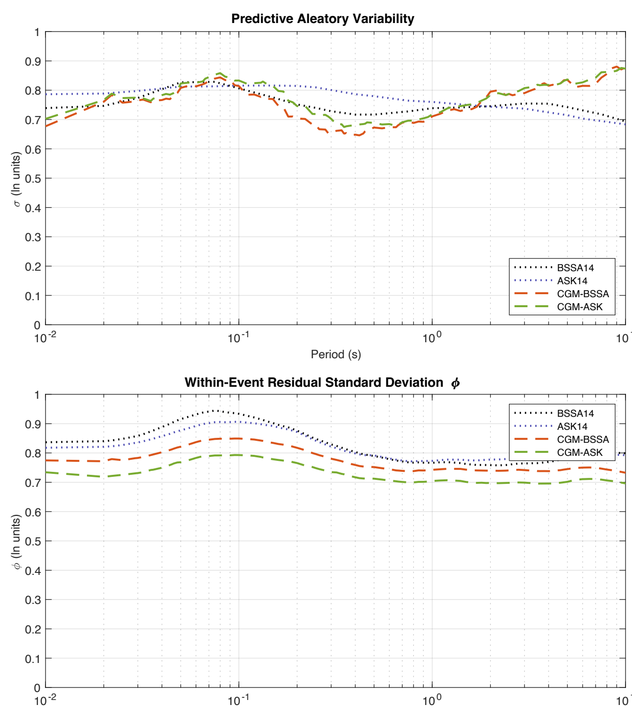

Figure 3: (top) Mean predictive σ (scaled std of 100 CVAE samples) vs. GMPE analytical σ, averaged over all records., (bottom) σ of residuals (obs $-$ GMPE median) vs. period for BSSA14, ASK14, CGM-BSSA, and CGM-ASK.

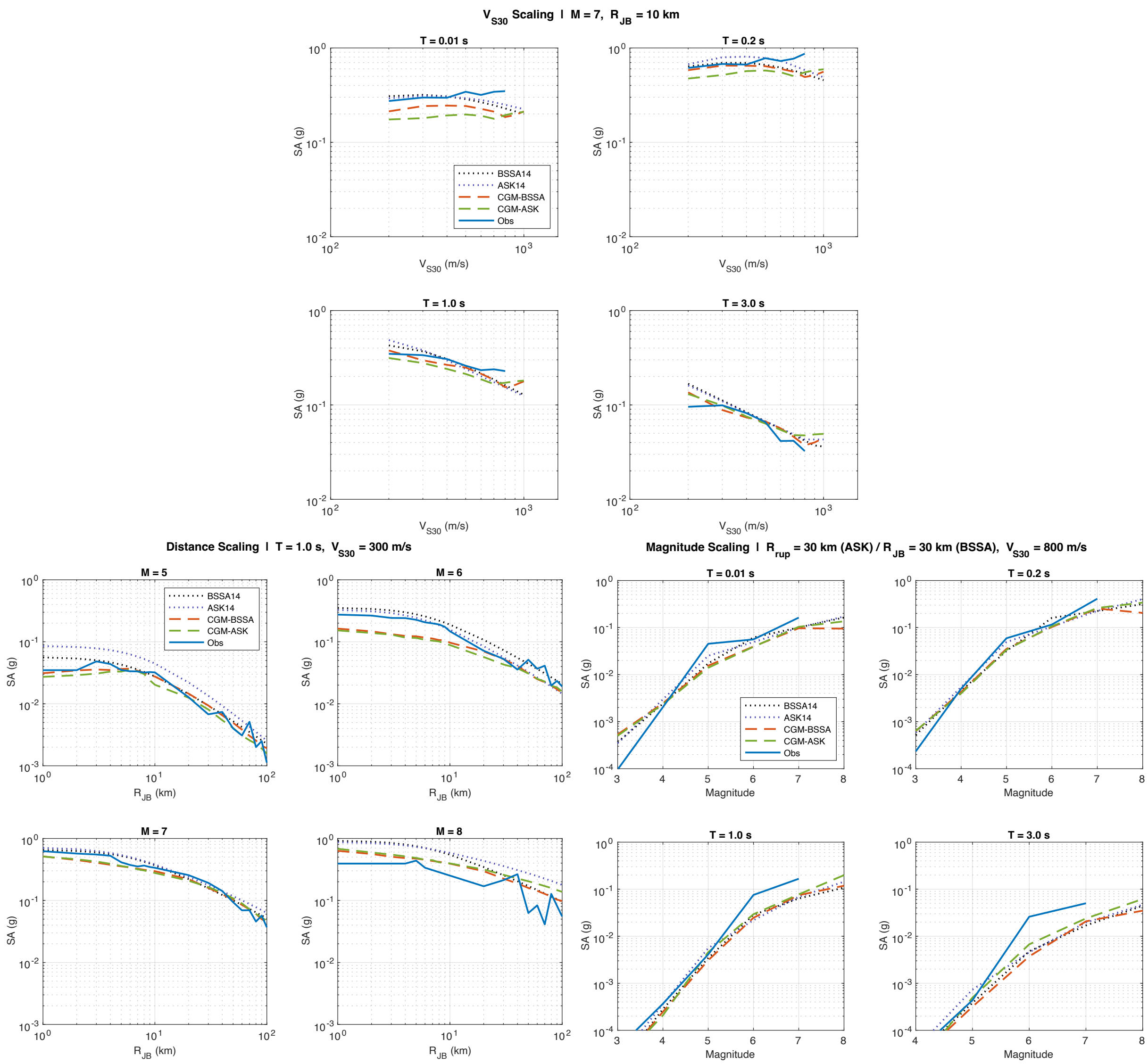

Figure 4: Scaling: (top) Median SA vs. V_{S30} at periods of 0.01, 0.2, 1.0 and 3.0 sec, (bottom-left) Distance scaling at M=5,6,7 and 8. (bottom-right) Magnitude scaling at periods of 0.01, 0.2, 1.0 and 3.0 sec.

References

N. A. Abrahamson, W. J. Silva, and R. Kamai. Summary of the ASK14 Ground Motion Relation for Active Crustal Regions. *Earthquake Spectra*, 30(3):1025–1055, 2014. ISSN 8755-2930. doi: 10.1193/070913eqs198m.

T. D. Ancheta, R. B. Darragh, J. P. Stewart, E. Seyhan, W. J. Silva, B. S. J. Chiou, K. E. Wooddell, R. W. Graves, A. R. Kottke, D. M. Boore, T. Kishida, and J. L. Donahue. NGA-West 2 Database. *Earthquake Spectra*, 30(3):989–1005, 01 2014. ISSN 8755-2930. doi: 10.1193/070913eqs197m.

D. M. Boore, J. P. Stewart, E. Seyhan, and G. M. Atkinson. NGA-West2 Equations for Predicting PGA, PGV, and 5% Damped PSA for Shallow Crustal Earthquakes. *Earthquake Spectra*, 30(3): 1057–1085, 2013. ISSN 8755-2930. doi: 10.1193/070113eqs184m.

M. A. Florez, M. Caporale, P. Buabthong, Z. E. Ross, D. Asimaki, and M.-A. Meier. Data-driven synthesis of broadband earthquake ground motions using artificial intelligence. *Bulletin of the Seismological Society of America*, 112(4):1979–1996, 2022.

N. Gregor, N. A. Abrahamson, G. M. Atkinson, D. M. Boore, Y. Bozorgnia, K. W. Campbell, B. S.-J. Chiou, I. Idriss, R. Kamai, E. Seyhan, W. Silva, J. P. Stewart, and R. Youngs. Comparison of NGA-West2 GMPEs. *Earthquake Spectra*, 30(3):1179–1197, 2014. ISSN 8755-2930. doi: 10.1193/070113eqs186m.

P. Ren, R. Nakata, M. Lacour, I. Naiman, N. Nakata, J. Song, Z. Bi, O. A. Malik, D. Morozov, O. Azencot, et al. Learning physics for unveiling hidden earthquake ground motions via conditional generative modeling. *arXiv preprint arXiv:2407.15089*, 2024.

Y. Shi, G. Lavrentiadis, D. Asimaki, Z. E. Ross, and K. Azizzadenesheli. Broadband ground-motion synthesis via generative adversarial neural operators: Development and validation. *Bulletin of the Seismological Society of America*, 114(4):2151–2171, 2024.

T. Wang, D. Trugman, and Y. Lin. Seismogen: Seismic waveform synthesis using gan with application to seismic data augmentation. *Journal of Geophysical Research: Solid Earth*, 126(4): e2020JB020077, 2021.