Canadian Conference - Pacific Conference on Earthquake Engineering 2023
Vancouver, British Columbia
June 25th – June 30th, 2023

Regional Ground Motion Generation using Gaussian Process Regression and Autoencoder

Chunxiao Ning[1] and Yazhou Xie[2*]

[1] PhD Student, Department of Civil Engineering, McGill University, Montreal, Quebec, Canada
[2] Assistant Professor, Department of Civil Engineering, McGill University, Montreal, Quebec, Canada
[*]tim.xie@mcgill.ca (Corresponding Author)

ABSTRACT

Rapid post-earthquake seismic response and damage assessment of structure portfolios at the regional scale are essential to increase situational awareness and help decision-makers plan risk-informed emergency responses. Such earthquake impact assessment calls for the availability of high-resolution seismic inputs at every structure location. To this end, this study leverages recorded motions and machine learning techniques to directly simulate event-based ground motion (GM) time series at target sites. Due to sparse sensor instrumentation, GMs are not recorded to cover all structure sites, yet they bear the potential to be utilized to predict seismic inputs at any locations of interest. To achieve this, the study employs the Gaussian process regression (GPR) and autoencoder (AE) to simulate GM time series at target sites based on the existing motion inventory recorded from one earthquake event. The AE is used to extract the recorded GMs' low-dimensional latent features, which are further treated as random Gaussian variables. The GPR model is then trained to fit the spatial distribution of these low-dimensional latent feature vectors, and the target GMs are generated by further decoding the latent features. The method is verified using the 2011 Tohoku earthquake GMs recorded by the seismograph network stations in Japan. The proposed method overcomes the long-standing difficulty of determining intra-sensor GMs and provides valuable motion inputs to facilitate regional seismic damage assessment.

Keywords: Ground motion simulation, Gaussian process regression, Autoencoder, Machine learning, Regional seismic damage assessment.

INTRODUCTION

A reliable assessment of structural damage after an earthquake event is essential for decision-makers to properly allocate resources for emergency responses [1]. Simulation-based seismic analyses of buildings and infrastructures need the inputs of earthquake ground motions (GMs), which can be obtained through motions recorded at spatially distributed seismograph stations. The distance between adjacent stations is generally larger than 10 km [2], resulting in that only a limited number of GM records are available. In addition, seismic damage of structure inventories within the intra-station space shows significant variability due to differences associated with geological conditions, structural geometry, material properties, design details, etc [3]. Therefore, interpolating intra-station GM time series is necessary for assessing the seismic damage of structures at specific locations of interest.

Earthquake GMs can be generated through the physics-based method, coherence function-based method, and data-driven method. Physics-based method [4] makes assumptions about the earthquake source, propagation media, and local site effects. The GMs are generated by utilizing physical laws to simulate the evolution of fault rupture and wave propagation in the nonlinear media. However, this method requires numerous underground shear-wave velocity data and fault parameters as the simulation inputs. The complete simulation process is also time-consuming, making it not applicable to achieve rapid post-earthquake damage assessment. The coherence function-based method employs the power spectra density model and coherence functions to construct the cross-power spectral matrix, which captures the spatial correlation of GMs [2,5]. The performance of the method is limited by the space and location of seismograph stations: it is more suitable for analyzing array records with small spacing. The data-driven method gains popularity in recent years because of the development of data acquisition, storage and processing techniques [6]. Algorithms such as inverse-distance-weighted (IDW) and kriging analysis have been utilized in previous studies to interpolate GMs based on recorded data. For example, the IDW method was used to estimate the peak ground acceleration (PGA) of the 2016 Kumamoto earthquake [7], while the accuracy of the method has not been presented in

detail. The interpolation of low-frequency GMs has been conducted through the array completion method [8], but this method requires the distance between stations to be within 1 km. Likewise, Huang and Wang [9] used wavelet packet decomposition and Kriging analysis-based spatial correlation of wavelet packet parameters to construct a record time history at an unmeasured site. Lu et al. [2] applied IDW to interpolate the response spectra at target locations, and used continuous wavelet transform to derive GM time series. These methods can be used to quickly obtain the GM fields in the disaster area, but their accuracy and applicability are questionable when 1) the station space is large; 2) GMs feature a variety of distinct features (e.g., frequency contents, magnitudes, durations).

To this end, this study develops a data-driven framework to generate regional GMs using recorded signals at seismograph stations. Convolutional variational autoencoder (CVAE) [10] first extracts the latent variables from high dimensional GM data. Afterwards, latent vectors are interpolated through the Gaussian process regression (GPR) and further converted back as GMs at target locations. The accuracy and applicability of the proposed framework are verified against the ground motion dataset recorded from the 2011 Tohoku earthquake.

VAE-GPR FRAMEWORK FOR REGIONAL GROUND MOTION GENERATION

As shown in Figure 1, the study proposes an end-to-end framework to interpolate spatially correlated GMs. The framework consists of three integrated parts: (1) signal transform and reconstruction of GMs through short-time Fourier transform (STFT); (2) dimension reduction and feature extraction through convolutional variational autoencoder (CVAE); (3) GM interpolation in latent space through Gaussian process regression (GPR). Technical details of each part are discussed below.

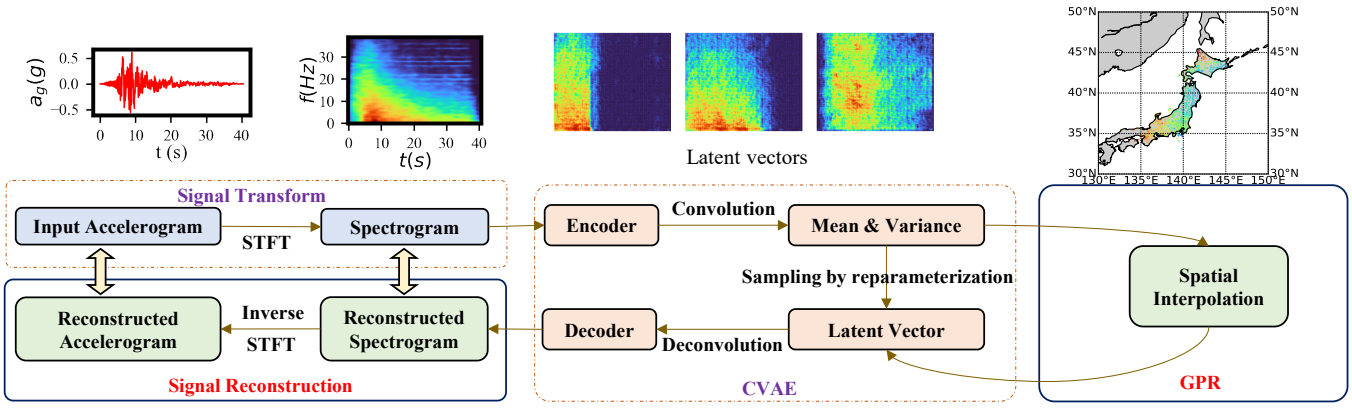

Figure 1. Proposed CNN surrogate model for structural response prediction.

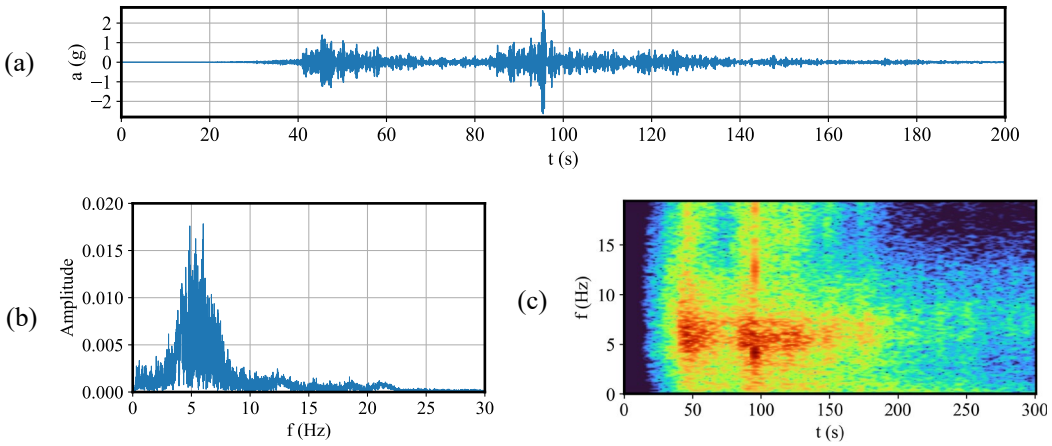

Figure 2. Ground motion representations: (a) waveform, (b) Fourier spectrum and (c) spectrogram.

Signal transformation and reconstruction through short-time Fourier transform

GM results from rupture movement at the earthquake epicenter and the propagation of the signal wave to different locations through distinct media. As such, the GMs recorded at spatially distributed seismograph stations show variations in frequency contents, amplitudes, and durations. GM is usually represented by a one-dimensional (1D) waveform in the time domain. However, the time domain representation does not expose the GM's frequency features, which are important elements for

nonstationary signals. In this regard, the spectrogram representation (Figure 2) is adopted here, given its ability to capture the temporal evolution of different frequency contents in time series [11]. The spectrogram can be developed by using various signal processing methods, such as STFT and wavelet analysis. This study adopts the STFT algorithm, as shown in Eq. (1), which subdivides the time series into small time windows with equal length, and performs Fourier transform on each individual time window.

$$S(\tau,\omega) = \int_{-\infty}^{\infty} x(t)w(t-\tau)e^{-i\omega t}dt \tag{1}$$

where $x(t)$ is the 1D time series, $w(\tau)$ is the window function, and $S(\tau,\omega)$ is the STFT value at frequency ω and time τ. It is noted that $S(\tau,\omega)$ is a complex number, with its spectrogram, namely the norm $|S(\tau,\omega)|$, captures the motion's energy distribution at different time instants and frequencies. Eq. (1) transforms 1D time series into 2D digital images, which facilitate the application of computer vision techniques, such as image classification [12,13], object detection [14], and semantic segmentation [15,16], etc. The STFT algorithm involves multiple hyperparameters, such as the length and overlapping of the time window, which may affect the time/frequency resolution of the spectrogram [11].

Reverting GM spectrograms to their original time series demands the STFT algorithm to be invertible. This reversibility is accomplished by utilizing window functions that conform to the constant overlap-add (COLA) constraint. [17]. The phase information for the waveform can be estimated iteratively through the Griffin-Lim algorithm [18]. As shown in Eq. (2), ISTFT uses spectrogram and phase information to compute the inverse Fourier transform at each time window, as well as overlaps and adds each individual inverted signal to the entire time sequence [19].

$$x(t) = \sum_{\tau=-\infty}^{\infty} \int_{-\frac{l}{2}}^{\frac{l}{2}} S(\tau,\omega)e^{i2\pi\omega\tau}d\omega \tag{2}$$

Convolutional variational autoencoder (CVAE)

The CVAE model is used in this study to reduce the dimensionality of the GM spectrogram and extract its latent features in a nonlinear manner [20,21]. As shown in Figure 3, the CVAE model is composed of an encoder and a decoder. Having different kernel, stride, and channel sizes, the convolutional layers [22] serve as the architecture for the encoder to compress the input spectrogram into a latent space and dig out its embedded latent features. In particular, model nonlinearity is introduced through the *ReLU* layer [23] that follows each convolutional layer. In contrast, the latent space is generated in a fully connected layer located at the end of the encoder. Subsequently, deconvolutional layers [24] are established to serve as the decoder for reconstructing the spectrogram from its latent features. Essentially, the deconvolutional layer can be regarded as a layer that combines the up-sampling layer and the 2D convolutional layer (*Conv2D*) [25].

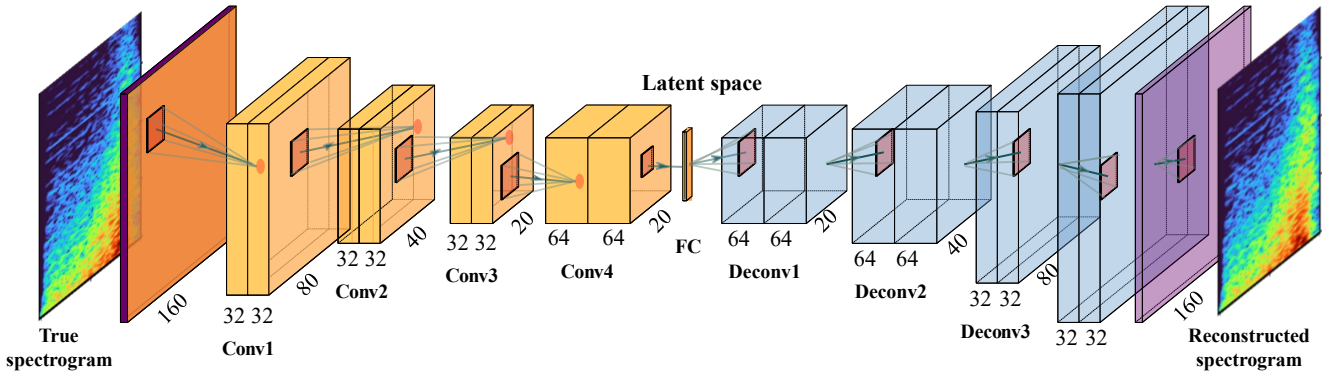

Figure 3. The architecture of the CVAE model.

The CVAE model encodes the input data into standard Gaussian distributions, through which the latent space is regularized and the model is less likely to overfit. As a result, data outputs from the encoder are the means and variances determining the distributions of latent variables; latent features can then be resampled using the reparameterization trick [26], as shown in Eq. (3).

$$z_i = \mu_i + \sigma_i \cdot \epsilon \tag{3}$$

where μ_i and σ_i are the mean and standard deviation for the i^{th} latent variable z_i, and ϵ is a noise term following a standard normal distribution.

The training of a CVAE model is to minimize two types of losses – the reconstruction loss between data inputs of the encoder and data outputs from the decoder, and the Kullback-Leibler (KL) divergence [27] that measures the distance between the latent distribution and standard Gaussian distribution. To further improve the reconstruction quality of the decoder, a discriminator [28] is trained to match the features of reconstructed spectrograms with those of true spectrograms, through which the reconstruction loss is defined as the difference between intermediate outputs from the discriminator. In a probabilistic representation, the latent variable z associated with GM spectrogram x is assumed to follow a prior standard Gaussian distribution $\mathrm{p}_\theta(z)$, whereas the developed CVAE model encodes its posterior distribution $\mathrm{q}_\phi(z|x)$ and decodes its distribution likelihood $p_\theta(x|z)$ for spectrogram reconstruction. To this end, the loss function for CVAE is formulated as Eq. (4).

$$L(\phi,\theta) = -\mathbb{E}_{q_\phi(z|x)}[log\ p_\theta(x|z)] + \beta KL\big(q_\phi(z|x)\big|\big|p_\theta(z)\big) \tag{4}$$

where the first term computes the reconstruction error, the second term is the KL divergence as a regularizer for the latent space, and the β value controls the strength of regularization. The encoder parameter ϕ and decoder parameter θ are learned during the training process to minimize the loss function in Eq. (4).

The CVAE model and discriminator are trained in an adversarial manner. Namely, the CVAE model aims to minimize the KL divergence and reconstruction loss (i.e., cheat the discriminator), while the discriminator aims to distinguish the reconstruction from the ground truth. Furthermore, the adaptive learning-rate method is considered to avoid neither slow convergence nor fast oscillation [29]. This study employs the adaptive moment estimation algorithm (Adam) as the adaptive learning method, given its efficiency in training the model with a large dataset and parameters, and its low storage memory requirement [30].

Gaussian process regression (GPR)

Gaussian process (GP) [31] is a non-parameterized supervised learning method that predicts the distribution of functions. The model assumes that every finite collection of data points from the latent function follows a multivariate normal distribution, whose correlation is determined by kernel functions. The data generation process from a GP model can be represented by Eq. (5):

$$f(\boldsymbol{x}) \sim \mathcal{GP}\big(m(\boldsymbol{x}), k(\boldsymbol{x},\boldsymbol{x}')\big) \tag{5}$$

where $m(\boldsymbol{x})$ is the mean function, and $k(\boldsymbol{x},\boldsymbol{x}')$ represents the kernel function. The choice of kernel functions depends on the expected pattern in the data, such as the smoothness. One of the most widely used kernel functions is the radial basis function (RBF) kernel, which assumes the correlation between two points decays exponentially with the increase of their distance, as formulated in Eq. (6). The σ_f and l are signal standard deviation and length scale, respectively. Matern kernels in Eq. (7) can be regarded as the generalization of the RBF kernel, with an additional parameter ν controlling the smoothness of the resultant function.

$$k_{RBF}(\boldsymbol{x},\boldsymbol{x}') = \sigma_f^2 \exp\left(-\frac{|\boldsymbol{x}-\boldsymbol{x}'|^2}{2l^2}\right) \tag{6}$$

$$k_{Matern}(\boldsymbol{x},\boldsymbol{x}') = \frac{1}{\Gamma(\nu)2^{\nu-1}}\left(\frac{\sqrt{2\nu}}{l}d(\boldsymbol{x},\boldsymbol{x}')\right)^{\nu} K_\nu\left(\frac{\sqrt{2\nu}}{l}d(\boldsymbol{x},\boldsymbol{x}')\right) \tag{7}$$

The latent variables of the recorded GMs at seismograph stations are used to fit the GPR model, which is capable of sampling the latent representation at any location. The prediction of intra-station GMs is achieved by decoding the interpolated latent variables using trained CVAE models.

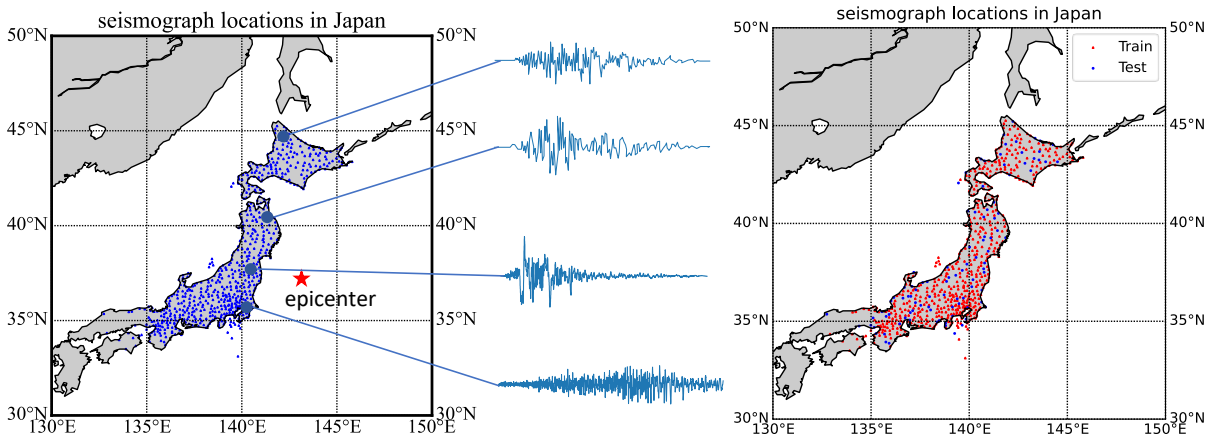

Figure 4. Earthquake GMs recorded at seismograph stations and split of the dataset.

CASE STUDY - 2011 TOHOKU EARTHQUAKE

The 2011 Tohoku earthquake was reported as the most powerful earthquake event ever recorded in Japan. The epicenter of the earthquake is located undersea in the Pacific Ocean, and the earthquake magnitude is 9.0 M_w. This megathrust earthquake and its subsequent tsunami induced 19,759 deaths, 6,242 injured, and 2,553 people missing [32]. After the event, the Japanese National Research Institute for Earth Science and Disaster Resilience (NIED) provided strong ground motion data collected by more than 1,000 seismograph stations nationwide. The locations of the seismograph stations are provided in Figure 4. In particular, 701 data samples are available from the NIED dataset for the 2011 Tohoku earthquake event. The selected data are shuffled and split into 601 training samples and 100 testing samples, as illustrated in Figure 4.

Data preprocessing and CVAE training

The selected datasets are transformed into spectrograms using STFT. The time window length for STFT is selected to be 2056 data points, and the overlapping length is considered 1/4 of the time window length. In addition, the Hanning function [33] is employed to avoid frequency leakage and satisfy the COLA requirement for STFT to be invertible. The time and frequency are truncated as 300s and 40Hz, respectively, without significant information loss.

The CVAE model is composed of an encoder with six convolutional layers, and a decoder with six deconvolutional layers [34]. The activation layers use the *ReLU* [35] function, while the output layer of the decoder implements the *sigmoid* function. The latent space is selected to have three latent variables, with which the CVAE model achieves the balance between reconstruction quality and latent space structurization. The discriminator has the same architecture as the encoder, with the output layer replaced by a logistic classifier.

The spectrograms are standardized before being fed into the CVAE model for training. Given the selected model architecture and hyperparameters, the CVAE model is trained against the 601 training samples using the mentioned mini-batch gradient descent algorithm with the Adam optimizer. The learning rate is set to be 1×10^{-4}, the batch size is selected as 50, and the training epoch is 3000. The trained model is capable of compressing spectrograms into low-dimensional representations and reconstructing them back by using the embedded data information in the latent space.

Figure 5 compares the reconstruction of a sample in the testing dataset with its ground truth. The spectrogram comparison indicates a good match of GM amplitudes at different time instants and frequency ranges. Essentially, the reconstructed spectrogram captures the duration characteristics of the GM. As proved through the time-history comparison in the same figure, the reconstructed accelerogram is on top of the original GM with the same PGA, phase information, frequency content, and duration. The mean values of the latent variables are standardized to the interval between 0 and 1, and the distributions of the three standardized latent variables are shown in Figure 6.

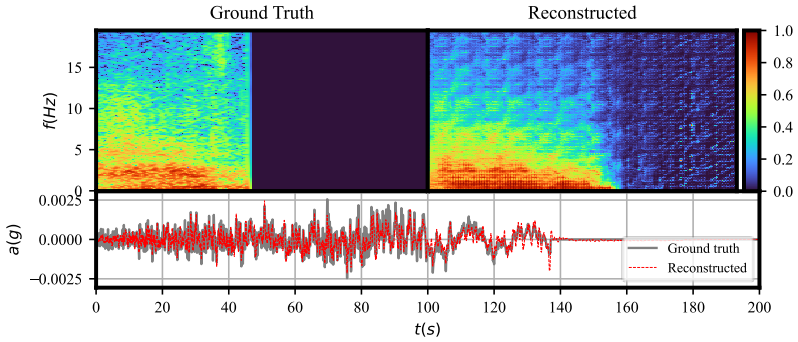

Figure 5. Comparison between CVAE reconstructed GM and ground truth.

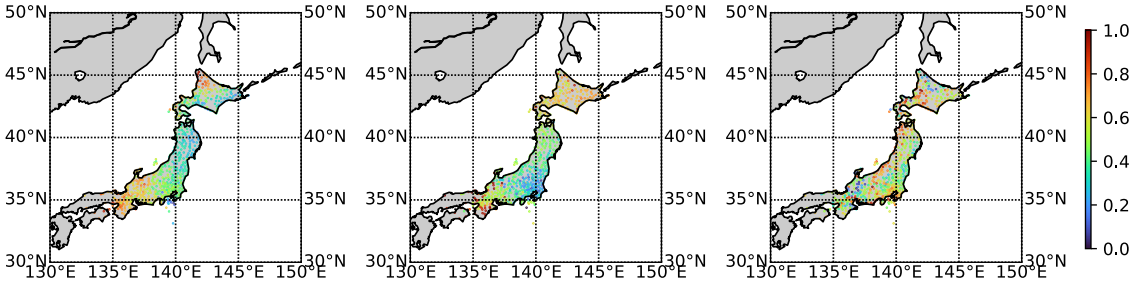

Figure 6. Distribution of latent variables.

GPR for GM interpolation

The CVAE model encodes the data samples in the training set into three latent variables, as shown in Figure 6. The GPR model is further used to fit those latent variables independently and predicts the associated spatial distributions. Latitude, longitude, and altitude values serve as the inputs for the GPR, and the latent variables constitute the output. The Matern kernels plus White noise kernels are used for GPR models, where the initial length scale and standard deviation are set to be 1.0 and 0.05, respectively. The L-BFGS-B algorithm [36] is used to optimize the GPR model, and the results are shown in Table 1. Figure 7 shows the spatial distributions of GPR-fitted latent variable mean values and the corresponding standard deviations. The latent codes shown in Table 1 are implicitly associated with GMs' physical features, such as duration, frequency content and source distance.

Table 1. Trained parameters for the GPR model.

Latent variables	Noise level	Length scale
Code 1	0.00622	1.070
Code 2	0.00935	0.985
Code 3	0.0230	0.617

The GPR model is further assessed by calculating the relative error between the GPR prediction and the true values for all the 701 samples in the dataset, as illustrated in Figure 7(d). It can be observed that the GPR model can accurately predict all three latent variables: most samples in both the training and testing sets have a relative error of less than 20%.

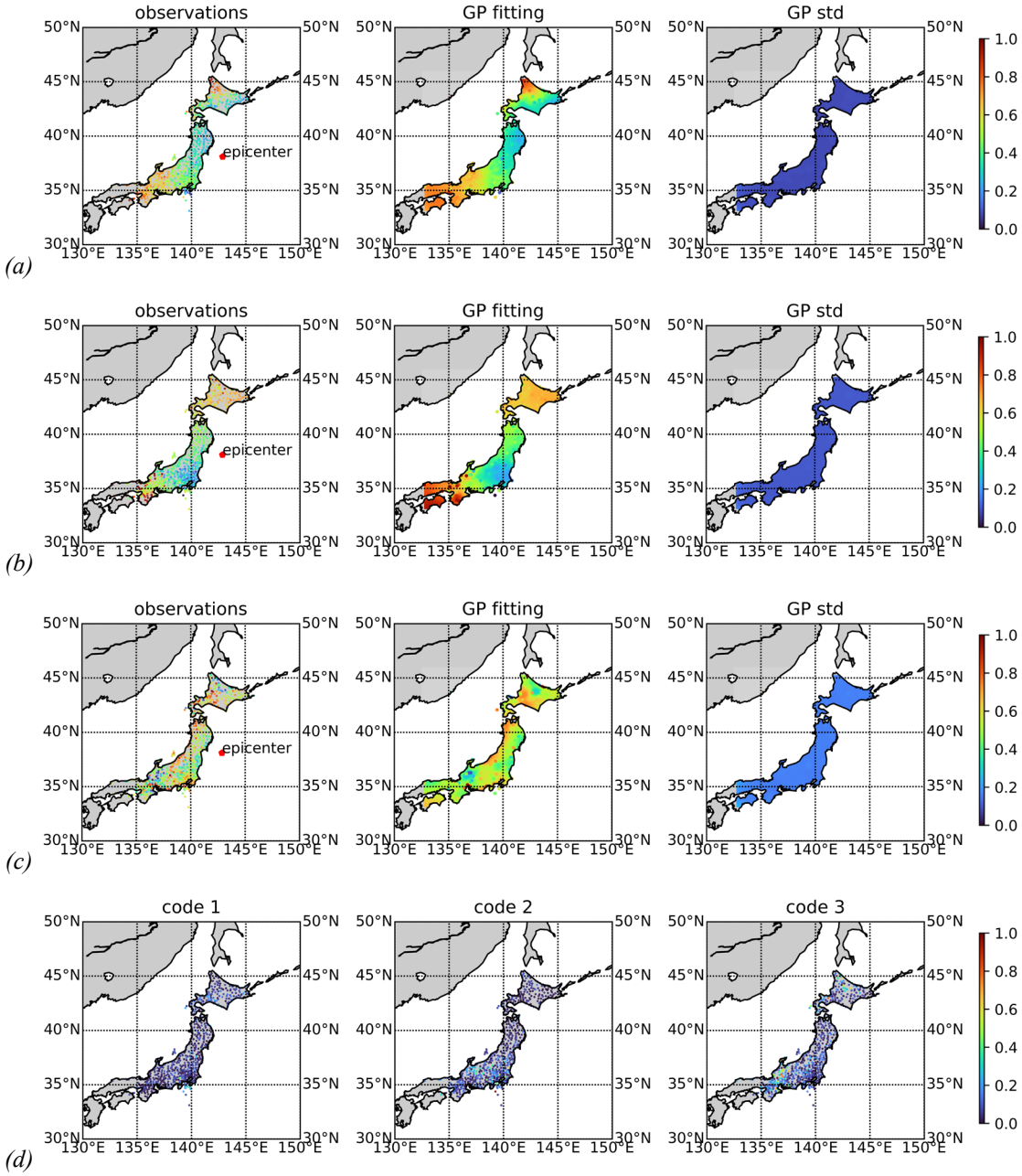

Figure 7. Distributions of latent codes fitted by GPR model. (a) latent varibale 1, (b) latent variable 2, (c) latent variable 3 and (d) relative error.

Decoding regional GMs

The GPR-interpolated latent variables are fed into the decoder for being reconstructed into GM spectrograms. The Griffin-Lim algorithm is used to find the phase information for ISTFT to estimate the time history of GMs. In addition, the GPR model provides a viable way to interpolate the peak ground accelerations (PGAs) for the region, and the reconstructed GM time series are scaled to the corresponding PGA at each location. Figure 8(a) shows the interpolation and standard deviation of PGA values for the earthquake event. The standard deviation is almost zeros everywhere, indicating the confident prediction of the GPR model. Figures 8(b) and 8(c) compare the predicted GMs in the training and testing sets with their ground truth data, respectively. The prediction is made by GPR interpolation and CVAE decoding. The comparison demonstrates that the proposed framework is capable of predicting GMs with accurate PGA, phase information, and duration.

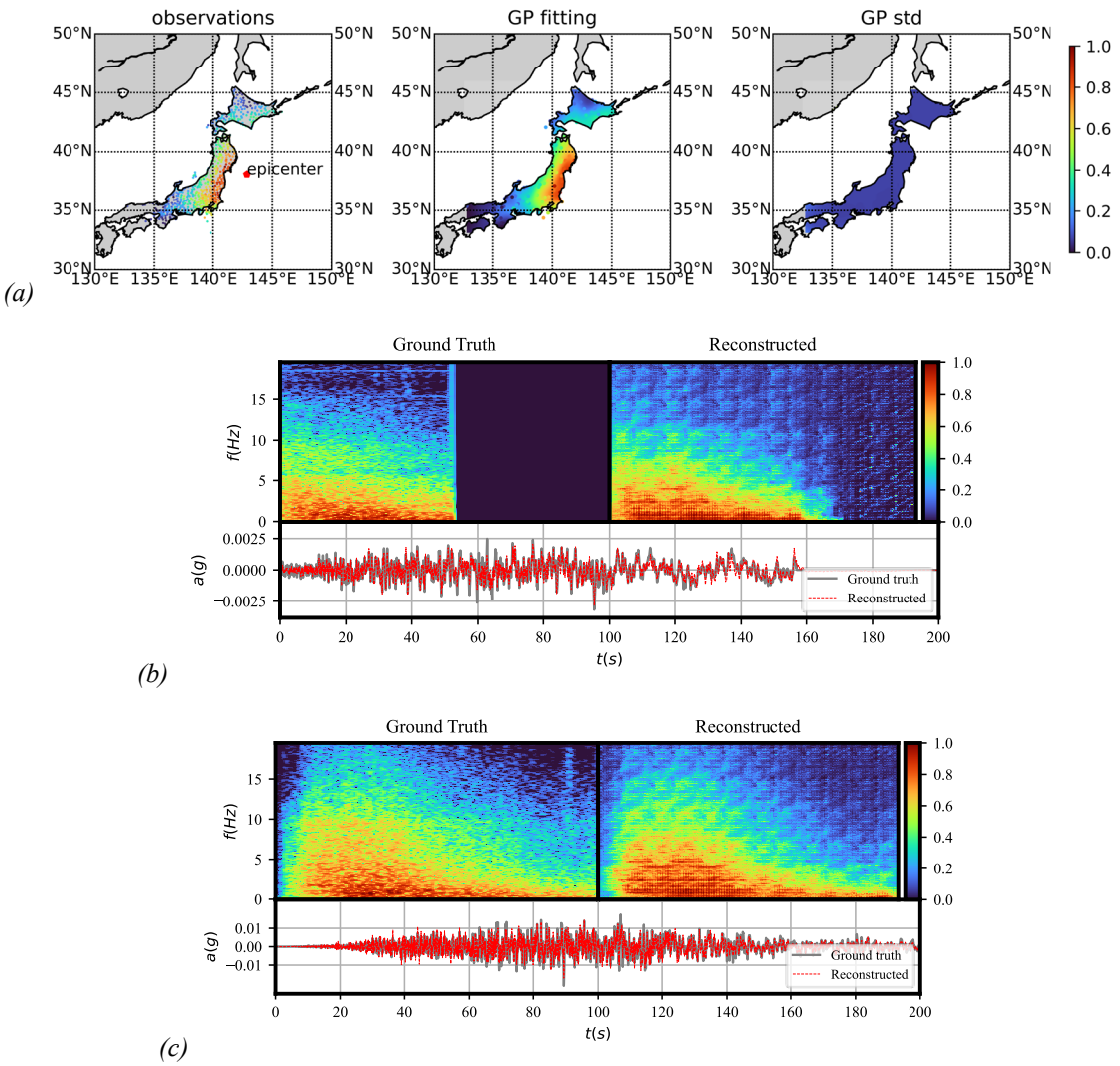

Figure 8. Interpolation of regional GMs. (a) Interpolation of peak ground acceleration, Comparison of a sample in the (b) training set, and (c) testing set.

CONCLUSIONS

The study integrates CVAE and GPR into an end-to-end analysis framework to predict the regional GM field based on a limited number of GMs recorded at seismograph stations. The STFT algorithm is first used to transform the GM time series into spectrograms, and the CVAE model compresses high-dimensional spectrograms into a low-dimensional latent space. The GPR models are trained to fit the spatial distribution of latent variables, whereas the trained CVAE model is further utilized to reconstruct latent variables into spectrograms. The ISTFT is finally used to transform spectrograms back to the time series data of GMs. The proposed framework is applied to the recorded GM data from the 2011 Tohoku earthquake event to predict intrastation GMs. The case study reveals that the predicted GMs capture the amplitude, frequency content, and duration characteristics of the ground truth data. The proposed framework provides valuable GM inputs to enable rapid post-earthquake seismic damage assessment at the regional scale.

The CVAE and GPR models in the study are trained independently, which cannot explicitly consider the dependency between the latent variable distribution and the spatial correlation of GMs. In future work, a joint CVAE-GPR model will be developed and optimized to capture the possible dependency between latent variables and spatial correlations spreading the region.

ACKNOWLEDGMENTS

This research has been supported by the NSERC Discovery Grant, Alliance Grant, and Mitacs Grant. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the official views or policies of the funding agencies.

REFERENCES

[1] Mackie KR, Wong J-M, Stojadinović B. Post-earthquake bridge repair cost and repair time estimation methodology. Earthq Eng Struct Dyn 2009:n/a-n/a. https://doi.org/10.1002/eqe.942.
[2] Lu X, Cheng Q, Tian Y, Huang Y. Regional Ground-Motion Simulation Using Recorded Ground Motions. Bull Seismol Soc Am 2021;111:825–38. https://doi.org/10.1785/0120200243.
[3] Ning C, Xie Y. Risk-Based Optimal Design of Seismic Protective Devices for a Multicomponent Bridge System Using Parameterized Annual Repair Cost Ratio. J Struct Eng 2022;148. https://doi.org/10.1061/(ASCE)ST.1943-541X.0003330.
[4] Atkinson GM, Assatourians K. Implementation and validation of EXSIM (a stochastic finite-fault ground-motion simulation algorithm) on the SCEC broadband platform. Seismol Res Lett 2015;86:48–60.
[5] Zerva A, Zervas V. Spatial variation of seismic ground motions: an overview. Appl Mech Rev 2002;55:271–97.
[6] Xie Y, Ebad Sichani M, Padgett JE, DesRoches R. The promise of implementing machine learning in earthquake engineering: A state-of-the-art review. Earthq Spectra 2020;36:1769–801. https://doi.org/10.1177/8755293020919419.
[7] Suzuki W, Aoi S, Kunugi T, Kubo H, Morikawa N, Nakamura H, et al. Strong motions observed by K-NET and KiK-net during the 2016 Kumamoto earthquake sequence. Earth, Planets Sp 2017;69:1–12.
[8] Thrainsson H. Modeling of earthquake ground motion in the frequency domain. Stanford University; 2000.
[9] Huang D, Wang G. Stochastic simulation of regionalized ground motions using wavelet packets and cokriging analysis. Earthq Eng Struct Dyn 2015;44:775–94.
[10] Doersch C. Tutorial on Variational Autoencoders 2016.
[11] Rabiner L, Schafer R. Theory and applications of digital speech processing. Prentice Hall Press; 2010.
[12] Wang N, Zhao Q, Li S, Zhao X, Zhao P. Damage Classification for Masonry Historic Structures Using Convolutional Neural Networks Based on Still Images. Comput Civ Infrastruct Eng 2018;33:1073–89. https://doi.org/10.1111/mice.12411.
[13] Liang X. Image-based post-disaster inspection of reinforced concrete bridge systems using deep learning with Bayesian optimization. Comput Civ Infrastruct Eng 2019;34:415–30. https://doi.org/10.1111/mice.12425.
[14] Shen J, Yan W, Li P, Xiong X. Deep learning-based object identification with instance segmentation and pseudo-LiDAR point cloud for work zone safety. Comput Civ Infrastruct Eng 2021;36:1549–67. https://doi.org/10.1111/mice.12749.
[15] Wang Z, Zhang Y, Mosalam KM, Gao Y, Huang S. Deep semantic segmentation for visual understanding on construction sites. Comput Civ Infrastruct Eng 2021:mice.12701. https://doi.org/10.1111/mice.12701.
[16] Zhang X, Rajan D, Story B. Concrete crack detection using context-aware deep semantic segmentation network. Comput Civ Infrastruct Eng 2019;34:951–71. https://doi.org/10.1111/mice.12477.
[17] Crochiere RE. A Weighted Overlap-Add Method of Short-Time Fourier Analysis/Synthesis. IEEE Trans Acoust Speech Signal Process 1980;28:99–102.
[18] Griffin D, Jae Lim. Signal estimation from modified short-time Fourier transform. IEEE Trans Acoust 1984;32:236–43. https://doi.org/10.1109/TASSP.1984.1164317.
[19] Oppenheim A V., Schafer RW, Buck JR. Discrete-Time Signal Processing. Prentice Hall 1999.
[20] Mousavi SM, Zhu W, Ellsworth W, Beroza G. Unsupervised Clustering of Seismic Signals Using Deep Convolutional Autoencoders. IEEE Geosci Remote Sens Lett 2019;16:1693–7. https://doi.org/10.1109/LGRS.2019.2909218.
[21] Ning C, Xie Y, Sun L. LSTM, WaveNet, and 2D CNN for nonlinear time history prediction of seismic responses. Eng Struct 2023;286:116083. https://doi.org/10.1016/j.engstruct.2023.116083.
[22] Schmidhuber JJ. Deep Learning in Neural Networks: An Overview. Neural Networks 2014;61:85–117. https://doi.org/10.1016/j.neunet.2014.09.003.
[23] Hannun A, Case C, Casper J, Catanzaro B, Diamos G, Elsen E, et al. Deep Speech: Scaling up end-to-end speech recognition 2014.
[24] Zeiler MD, Krishnan D, Taylor GW, Fergus R. Deconvolutional networks. 2010 IEEE Comput. Soc. Conf. Comput. Vis. pattern Recognit., IEEE; 2010, p. 2528–35.
[25] Chollet F. Keras: The python deep learning library. Astrophys Source Code Libr 2018:ascl-1806.
[26] Kingma DP, Welling M. Auto-Encoding Variational Bayes. 2nd Int Conf Learn Represent ICLR 2014 - Conf Track Proc 2013:1–14.
[27] Kullback S. Information theory and statistics. Courier Corporation; 1997.

[28] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks. Commun ACM 2020;63:139–44.
[29] Ruder S. An overview of gradient descent optimization algorithms 2016:1–14.
[30] Kingma DP, Ba J. Adam: A method for stochastic optimization. ArXiv Prepr ArXiv14126980 2014.
[31] Schulz E, Speekenbrink M, Krause A. A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. J Math Psychol 2018;85:1–16.
[32] Nakahara S, Ichikawa M. Mortality in the 2011 Tsunami in Japan. J Epidemiol 2013;23:70–3. https://doi.org/10.2188/jea.JE20120114.
[33] Braun S. Windows. Encycl. Vib., Elsevier; 2001, p. 1587–95. https://doi.org/10.1006/rwvb.2001.0052.
[34] Dumoulin V, Visin F. A guide to convolution arithmetic for deep learning 2016:1–28.
[35] Xu B, Wang N, Chen T, Li M. Empirical Evaluation of Rectified Activations in Convolutional Network 2015.
[36] Zhu C, Byrd RH, Lu P, Nocedal J. Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. ACM Trans Math Softw 1997;23:550–60.