ROSERS -Real-time On-Site Estimation of Response Spectra

Jawad Fayaz
Teesside University

12th National Conference
on Earthquake Engineering
Salt Lake City, Utah
27 June - 1 July 2022

Hosted by the Earthquake Engineering Research Institute

ROSERS - Real-time On-Site Estimation of Response Spectra

J. Fayaz[1] and C. Galasso[2]

ABSTRACT

Various earthquake early warning (EEW) methodologies have been proposed globally for speedily estimating information (i.e., location, magnitude, ground-shaking intensities, and/or potential consequences) about ongoing seismic events for real-time/near real-time earthquake risk management. Conventional EEW methods have often been based on the inferred physics of the fault ruptures combined with simplified empirical models to estimate the source parameters and intensity measures of interest. This approach can lead to significant uncertainties in real-time forecasting and decision-making. Given the recent boost in computational resources, data-driven methods/models are widely accepted as effective alternatives for EEW. This study introduces a highly accurate deep-learning-based computational framework named ROSERS (Real-time On-Site Estimation of Response Spectra) to estimate the acceleration response spectrum ($S_a(T)$) of the expected on-site ground motion waveforms using early non-damage-causing p-waves and site characteristics. The framework is trained using a carefully selected extensive database of ground motions. Due to the well-known correlation of $S_a(T)$ with structures' seismic response and resulting damage/losses, rapid and accurate knowledge of expected on-site $S_a(T)$ is highly beneficial to various end-users to make well-informed decisions. The framework is thoroughly assessed through multiple statistical tests and the analyses demonstrate that the overall framework leads to excellent prediction power to be implemented in a real-time setting.

Introduction

Earthquake early-warning (EEW) computational frameworks are generally developed to define different alert levels for life and asset protection from earthquake-induced ground shaking before reaching a given target site. Research on EEW is mainly oriented to efficiently provide accurate warnings and support risk-management decision-making during an earthquake event, *i.e.*, from the time of fault rupture to when the ground shaking ends at the target site. In a real-time setting, EEW utilizes information from the fast-traveling longitudinal *p*-waves (characterized by low-amplitude and generally non-damaging) of the arriving earthquake waveform to trigger alerts for preparing against the high-amplitude destructive transverse *s*-waves. Typically, *p*-waves are 1.7 times faster than *s*-waves in any given material [1]. EEW systems exploit this time lag between the arrival of the *p*-waves and *s*-waves to predict parameters of interest for real-time decision-making purposes.

In the field of Earthquake Engineering and performance-based earthquake engineering, the acceleration response spectrum, $S_a(T)$, is a widely used IM that can effectively integrate the characteristics of the ground motion waveform (such as amplitude, frequency content, etc.) with the dynamic behavior of a structural system idealized as a single-degree-of-freedom system (SDOF) (e.g.,[2], [3]). An accurate estimation of the ground-motion $S_a(T)$ values in real-time can help stakeholders in risk-informed decision making [4]. This paper contributes extending the objectives of current EEW frameworks to be accurate, timely, and both structurally- and ground motion-informed. In particular, it proposes a highly accurate and novel on-site EEW framework named ROSERS (Real-time On-Site Estimation of Response Spectra) developed

[1]Research Fellow, Department of Civil, Environmental, and Geomatic Engineering, University College London, UK WC1E 6BT (email: j.fayaz@ucl.ac.uk)

[2]Professor, Department of Civil, Environmental, and Geomatic Engineering, University College London, UK WC1E 6BT

Fayaz J., Galasso C. ROSERS - Real-time On-Site Estimation of Response Spectra. *Proceedings of the 12th National Conference in Earthquake Engineering*, Earthquake Engineering Research Institute, Salt Lake City, UT. 2022.

using deep neural network (DNN) and variational autoencoder (VAE). The framework uses site characteristics of the recording station and an IM vector (representing amplitude, energy, duration, and frequency content) of the initial three seconds of the arriving ground motion (after detection of *p*-waves) to estimate a vector of PGA (PGA is also referred to $S_a(T = 0)$) and 95-period $S_a(T)$ spectrum of the expected on-site complete ground motion waveform. Hence, apart from utilizing a more structurally informed IM as a target decision variable for EEW, this study utilizes an advanced neural-network approach to optimally conduct dimensionality reduction of the $S_a(T)$ spectrum into surrogate latent variables that can be practically used in real-time. Furthermore, using data-driven and deep-learning-based methods minimizes the assumptions involved in the model development and makes the framework flexible and easily re-trainable. In addition, the high dimensional interpolative and extrapolative nature of DNNs makes this study more robust than non-deep learning-based and traditional methods [5], [12]. The proposed framework is observed to perform well on the used dataset, and based on the findings of this paper, it can be highly beneficial to advance EEW research and support various stakeholders.

Proposed Framework

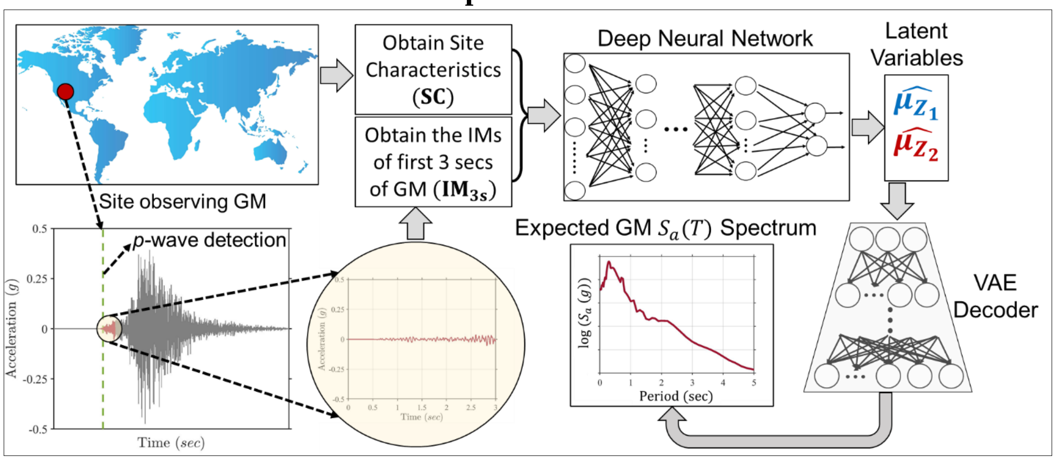


Figure 1. Illustration of the ROSERS framework. The illustration flows from left to right. The implementation of this ROSERS framework, on average, takes ~0.75 seconds on a modern personal computing machine.

The proposed ROSERS framework is illustrated in Figure 1 [18]. A digital recording station continuously monitors ground shaking at a given location. As the seismic sensor detects the initial *p*-waves of the arriving ground motion ([6], [7]), the ROSERS framework stands by to receive three seconds of the incoming ground motion. The framework utilizes the first three seconds of the observed ground motion in real-time to compute a vector of seven intensity measures denoted as $\mathbf{IM_{3s}}$. The $\mathbf{IM_{3s}}$ correspond to the amplitude, energy, frequency, and significant duration of the initial three seconds of the ground-motion waveform after detection of *p*-waves. The computed $\mathbf{IM_{3s}}$ are combined with the vector of site characteristics (denoted as **SC**) of the recording station, known *apriori*. The obtained $\mathbf{IM_{3s}}$ and **SC** are then used as inputs to a pre-trained feed forward DNN that estimates two mean latent variables ($\widehat{\mu_{z_1}}$ and $\widehat{\mu_{z_2}}$) which are then cascaded to a pre-trained decoder of VAE. It should be noted here that the two latent variables are not physically derived but represent two-dimensional statistical surrogacy of the complete $S_a(T)$ spectrum, which is obtained by training VAE. The VAE decoder transforms the estimated latent variables ($\widehat{\mu_{z_1}}$ and $\widehat{\mu_{z_2}}$) into the vector of PGA and 95-period $S_a(T)$ spectrum (denoted as $\mathbf{S_a}$). The obtained $\mathbf{S_a}$ vector can be utilized by stakeholders to informatively alert a community through risk informed EEW decision support systems, develop shakemaps, perform structural control, and other similar applications in real-time. Following sections describe various components of the proposed framework.

Ground Motion Database

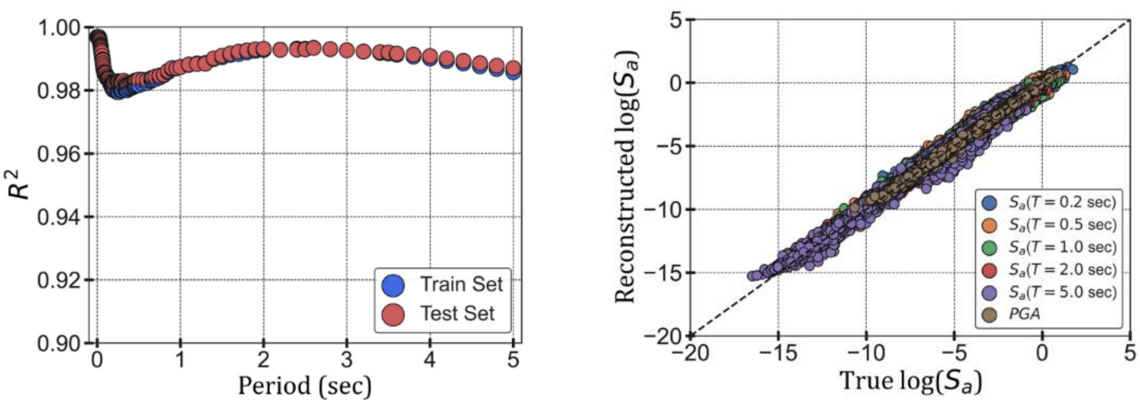
The framework is developed and trained using the recorded ground motions available in the comprehensive database of the Next-Generation Attenuation West version 2 (NGA-West2) project [8]. The database utilized in this study consists of 13916 ground motion components from 277 seismic events recorded worldwide and are processed based on the recommendations of [9]. It is worth noting that ground motions in real-time are generally not processed in the same manner, and the processing usually involves the use of simplistic techniques such as trend removals and low-pass and high-pass filtering [10], [11]. However, this assumption is not expected to significantly change the proposed framework's predictive performance; the data-driven nature of the artificial neural networks, with sufficient training, can quickly make the framework adapt to any minor deviances and noise ([5], [12]). It was observed that the database is skewed towards events with $M \leq 5.5$. Hence, in this study, the ground motions with $M \leq 5.5$ (majority class) are undersampled such that the number of ground motions with $M \leq 5.5$ and $M > 5.5$ is equivalent while maintaining the distributions of R_{rup} and site's average shear-wave velocity of the soil up to 30m (V_{s30}). This results in 6392 ground motion components whose *p*-waves arrival times are then deduced using an automated *p*-phase picker (PPHASEPICKER) algorithm developed by Kalkan (2016) [6]. Only the ground motion time histories after *p*-wave arrival are considered for the computation of $\mathbf{IM_{3s}}$ and further analysis in this study.

Components of the ROSERS Framework

Variational Autoencoder (VAE)

The $\mathbf{S_a}$ vector is computed for the 6392 ground motion components for the 96 periods and used as inputs to be reconstructed in the VAE [13]. The VAE is bottlenecked to have two independent normally distributed latent variables (denoted as z_1 and z_2 with means μ_{z_1} and μ_{z_2} and variances $\sigma_{z_1}^2$ and $\sigma_{z_2}^2$, respectively) in the sampling layer. A two-dimensional latent variable space is used as it results in a good trade-off between higher reconstruction power and a lower number of latent dimensions. The neural network-based VAE is trained through cross-validation using a randomly split 80% of the dataset while the remaining 20% is used as the final test set. The training and testing of the VAE is performed using a log transformation of $S_a(T)$. The VAE provides a probabilistic approach to describe a vectorial observation in their latent variable space. This is done using a neural network-based encoder (recognition model) trained in conjunction with a neural network-based decoder (generative model) that can use the latent variable space to reconstruct the observations. This means that the encoder describes a probability distribution for each latent attribute from which values are randomly sampled to be fed into the decoder that is expected to accurately reconstruct the input. Hence, the latent variable space is essentially compelled to possess continuous and smooth representations. Consequently, values nearby to one another in the latent space correspond to similar reconstructions using the decoder.

The coefficient of determination R^2 between the true $\mathbf{S_a}$ and reconstructed $\widehat{\mathbf{S_a}}$ for the 96 periods for both training and test sets is observed to be above 0.98, demonstrating excellent reconstruction power of the developed VAE with minimal bias and variance. This can be further observed from Figure 2b, where the true *vs.* reconstructed $S_a(T)$ are presented for periods 0.2 sec, 0.5 sec, 1 sec, 2 sec, 5 sec and 0 sec (PGA). Based on Figures 2a and 2b, it is clear that the latent variables z_1 and z_2 can adequately reconstruct the $\mathbf{S_a}$ using the VAE-decoder. Also, the utilization of the surrogate latent variables and dimensionality reduction makes sure that the estimated $\mathbf{S_a}$ is inherently cross correlated in prediction



(a) (b)

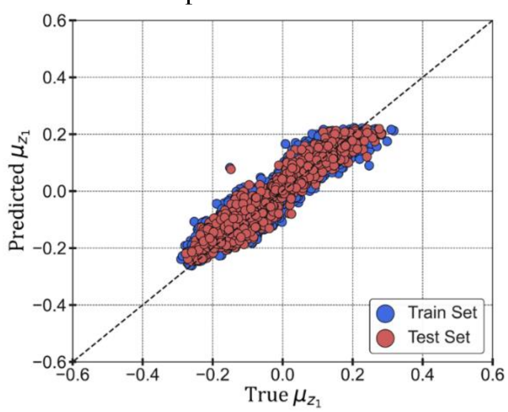
Figure 2. (a) R^2 of reconstructed $S_a(T)$ at various periods; and (b) True *vs.* reconstructed $S_a(T)$

Estimation of Latent Variables using Deep Neural Networks (DNNs)

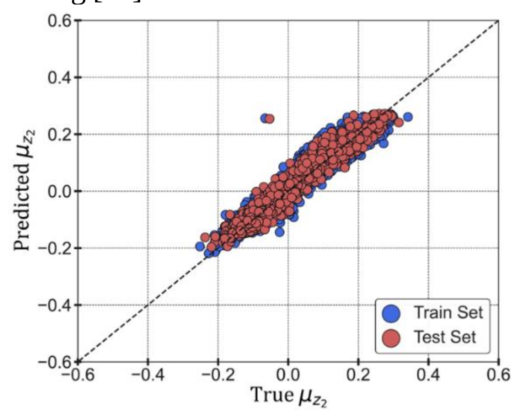
In a real-time setting, to construct the S_a of the expected ground motion, it is essential to accurately estimate the latent variables z_1 and z_2 which can then be used in the VAE decoder. To allow ample on-site warning time during the occurring ground motion, which can last up to 1 to 2 minutes ([14], [15]), early three seconds of the arriving waveform is used for predictions and early warning. The three seconds of waveform is used to compute $\mathbf{IM_{3s}}$ vector which includes: Arias

Intensity (I_a) (in *m/s*), Significant Duration (D_{5-95}) (in *sec*), Mean Period (T_m) (in *sec*), Peak Ground Acceleration (PGA) (in *g*), Peak Ground Velocity (PGV) (in *m/s*), Peak Ground Displacement (PGD) (in *m*), and Cumulative Absolute Velocity (CAV) (in *m/s*). Site characteristics (**SC**) are quantified using the site shear-wave velocity (V_{s30}) and basin depth to a shear velocity of 2.5 km/s ($Z_{2.5}$).

To estimate the μ_{z_1} and μ_{z_2} using the **SC** and $\mathbf{IM_{3s}}$ vectors, a 15 layered DNN with two nodes and linear activation function in the output layer is trained. The DNN is trained using a training dataset (randomly selected as 80% of the entire dataset), while evaluations are conducted on the remaining test dataset. The training is conducted using stochastic gradient descent with dropout and early stopping regularizations[16]. The true *vs.* predicted μ_{z_1} and μ_{z_2} from the trained DNN are presented in Figures 3a and 3b, respectively. The nature of predictions is observed to be very close to 1:1 for both train and test sets, thereby indicating the good prediction power of the DNN to estimate the two latent variables. The mean predictions of the trained DNN led to R^2 of 0.91 and 0.93 for μ_{z_1} and μ_{z_2}, respectively. Apart from high prediction power, another advantage of using DNNs is that both the latent variables are estimated simultaneously. A correlation coefficient of 0.91 was observed in DNN's predictions of μ_{z_1} and μ_{z_2} as compared to the true correlation coefficient of 0.88. This means the trained DNNs are highly successful in ensuring that the estimations are implicitly linked, *i.e.*, properly reflecting their internal correlation. Furthermore, due to the hierarchical nature of the data (*i.e.*, multiple recordings from the same event), the residuals between the $\mathbf{S_a}_{ij}$ and predictions of VAE-DNN $\widehat{\mathbf{S_a}}_{ij}$ are used to develop inter-event (**Φ**) and within-event (**T**) covariance matrices for the 96 periods for the i^{th} event and j^{th} recording [17].



(a)



(b)

Figure 3. True *vs*. predicted (a) μ_{z_1} and (b) μ_{z_2} from the trained DNN

Conclusions

This study introduces a deep learning-based on-site earthquake early warning (EEW) framework named ROSERS (Real-time On-Site Estimation of Response Spectra). The framework utilizes a pre-trained deep neural network (DNN) to estimate two latent variables (that represent statistical surrogates of the $S_a(T)$ spectrum) using the site characteristics and initial three seconds of on-site ground motion after *p*-wave detection. The two latent variables are then used in a trained decoder of variational autoencoder (VAE) to estimate peak ground acceleration (PGA) and 95-period spectral acceleration ($S_a(T)$) response spectrum for the expected complete on-site ground motion. While many other research studies have tried to use machine learning and deep learning frameworks for EEW purposes, they mainly concentrate on predicting ground acceleration, which can be significantly different from the acceleration caused in the structures (dynamic amplification). The prediction of the complete $S_a(T)$ spectrum (while inherently maintaining the cross-correlations) with only three seconds of the initial ground motion waveform in a highly accurate manner coupled with an average prediction time of less than one second makes this research unique and can contribute to the improvement of the current EEW decision-making and near-real-time rapid response systems.

Acknowledgments

This research is funded by the European Union's Horizon 2020 research and innovation program, specifically grant agreement number 821046: TURNkey "Towards more Earthquake-resilient Urban Societies through a Multi-sensor-based Information System enabling Earthquake Forecasting, Early Warning and Rapid Response actions."

References

1 S. L. Kramer, *Geotechnical Earthquake Engineering*. Prentice Hall, 1996.

2 P. Bazzurro, C. A. Cornell, N. Shome, and J. E. Carballo, "Three Proposals for Characterizing MDOF Nonlinear Seismic Response," *Journal of Structural Engineering*, vol. 124, no. 11, pp. 1281–1289, 1998, doi: 10.1061/(asce)0733-9445(1998)124:11(1281).

3 D. Vamvatsikos and C. Allin Cornell, "Incremental dynamic analysis," *Earthquake Engineering and Structural Dynamics*, vol. 31, no. 3, pp. 491–514, 2002, doi: 10.1002/eqe.141.

4 G. Cremen and C. Galasso, "Earthquake early warning: Recent advances and perspectives," *Earth-Science Reviews*, vol. 205, no. April, p. 103184, 2020, doi: 10.1016/j.earscirev.2020.103184.

5 J. Jumper *et al.*, "Highly accurate protein structure prediction with AlphaFold," *Nature*, vol. 596, no. 7873, pp. 583–589, 2021, doi: 10.1038/s41586-021-03819-2.

6 E. Kalkan, "An automatic P-phase arrival-time picker," *Bulletin of the Seismological Society of America*, vol. 106, no. 3, pp. 971–986, 2016, doi: 10.1785/0120150111.

7 T. Akazawa, "A Technique for Automatic Detection of Onset Time of P- and S-Phases in Strong Motion Records," *13th World Conference on Earthquake Engineering*, no. 786, p. 786, 2004.

8 T. D. Ancheta *et al.*, "NGA-West2 Database," *Earthquake Spectra*, vol. 30, no. 3, pp. 989–1005, Aug. 2014, doi: 10.1193/070913EQS197M.

9 D. M. Boore and J. J. Bommer, "Processing of strong-motion accelerograms: Needs, options and consequences," *Soil Dynamics and Earthquake Engineering*, vol. 25, no. 2, pp. 93–115, 2005, doi: 10.1016/j.soildyn.2004.10.007.

10 A. Caruso, S. Colombelli, L. Elia, M. Picozzi, and A. Zollo, "An on-site alert level early warning system for Italy," *Journal of Geophysical Research: Solid Earth*, vol. 122, no. 3, pp. 2106–2118, 2017, doi: 10.1002/2016JB013403.

11 T. Y. Hsu *et al.*, "Rapid on-site peak ground acceleration estimation based on support vector regression and P-wave features in Taiwan," *Soil Dynamics and Earthquake Engineering*, vol. 49, pp. 210–217, 2013, doi: 10.1016/j.soildyn.2013.03.001.

12 N. Borodinov, S. Neumayer, S. v. Kalinin, O. S. Ovchinnikova, R. K. Vasudevan, and S. Jesse, "Deep neural networks for understanding noisy data applied to physical property extraction in scanning probe microscopy," *npj Computational Materials*, vol. 5, no. 1, pp. 1–8, 2019, doi: 10.1038/s41524-019-0148-5.

13 D. P. Kingma and M. Welling, "An introduction to variational autoencoders," *Foundations and Trends in Machine Learning*, vol. 12, no. 4, pp. 307–392, 2019, doi: 10.1561/2200000056.

14 J. Fayaz, M. Medalla, and F. Zareian, "Sensitivity of the response of Box-Girder Seat-type bridges to the duration of ground motions arising from crustal and subduction earthquakes," *Engineering Structures*, vol. 219, no. April, p. 110845, 2020, doi: 10.1016/j.engstruct.2020.110845.

15 J. Fayaz, Y. Xiang, and F. Zareian, "Generalized ground motion prediction model using hybrid recurrent neural network," *Earthquake Engineering and Structural Dynamics*, vol. 50, no. 6, pp. 1539–1561, 2021, doi: 10.1002/eqe.3410.

16 F. Chollet, *Deep Learning with Python*. Manning, 2017.

17 E. Demidenko, *Mixed Models: Theory and Applications with R, 2nd Edition*, 2nd Edition. Wiley Publisher, 2013.

18 J. Fayaz and C. Galasso (2022). "A Deep Neural Network Framework for Real-Time On-Site Estimation of Acceleration Response Spectra of Seismic Ground Motions". *Computer-Aided Civil and Infrastructure Engineering*, 1-17, https://doi.org/10.1111/mice.12830