17th World Conference on Earthquake Engineering, 17WCEE
Sendai, Japan - September 13th to 18th 2020

Ground Motion Spectral Estimation using Recurrent Neural Network (RNN)

J. Fayaz[(1)], Y. Xiang[(2)], F. Zareian[(3)]

[(1)] *PhD Candidate, Department of Civil and Environmental Engineering, University of California, Irvine, CA, USA, jfayaz@uci.edu*
[(2)] *PhD Candidate, Department of Civil and Environmental Engineering, University of California, Irvine, CA, USA, yijunx3@uci.edu*
[(3)] *Associate Professor, Department of Civil and Environmental Engineering, University of California, Irvine, CA, USA, zareian@uci.edu*

Abstract

The currently used Ground Motion Prediction Models (GMPMs) are developed independently for different periods of a single-degree-of-freedom oscillator and do not account for internal correlations between the spectral accelerations at various periods for the same ground motion. The current approach to incorporating such a correlation is through a post-processing technique by Jayaram and Baker (2011); their method relates the spectral accelerations of different periods using linear correlation coefficients. The proposed correlation coefficients, however, are not highly exhaustive and can be improved to describe ground motion spectra in a more accurate manner. This study proposes a sophisticated Recurrent Neural Network (RNN) framework that can be used to develop ground motion spectra using similar inputs as GMMs. A series of statistical tests are conducted to assess the importance of various earthquake sources and site parameters in predicting spectral accelerations. The obtained parameters are then used as an input to the RNN-based ground motion spectral estimation framework. Contrary to previous research, the target of this framework is to estimate the entire spectrum instead of a single spectral acceleration. The discrepancy between the RNN-predicted spectrum and the true spectrum is minimized by maximizing the log-likelihood function using a non-convex optimization method Covariance Matrix Adaptation Evolution Strategy (CMA-ES). CMA-ES is used to obtain the inter-event covariance matrix and the intra-event covariance matrix to account for the hierarchical structure of the data. Hence, given the source and site parameters, the hybrid Recurrent Neural Network eventually returns an estimation of the GM spectrum along with an estimation of inter-event and intra-event errors. The RNN framework is finally validated by conducting a series of statistical tests on the residuals obtained from the model.

Keywords: Ground Motion Spectrum, Neural Networks, Spectral Correlations, Hazard

1. Introduction

In the field of structural and hazard analysis, ground motion prediction equations (GMPM) form an essential tool that is used to estimate the levels of ground motion intensity using the earthquake event parameters such as Magnitude M_w, Distance to Fault R_{rup}, etc. In general, GMPMs provide statistical predictions of peak ground acceleration (PGA), peak ground velocity (PGV), and pseudo-spectral acceleration (S_a) of a single-degree-of-freedom-system of various periods. GMPMs are used along with an Earthquake Rupture Forecast (ERF) database such as UCERF2 (2008), UCERF3 (2015) , which provide the GMPMs with the necessary event parameters to estimate ground motion intensity at the sites of interest. This is widely useful in an extensive range of research areas of Seismic Engineering such as regional seismic analysis, structural loss estimation, hazard mapping, etc. Over the years there have been numerous GMPMs developed across the world for both global and regional bases (Douglas 2019). Conventionally, GMPMs are developed as parametric functional forms using empirical data. Since the ground motions are recorded in two or three orthogonal directions, various measures have been used to quantify the intensity of ground motions on single-degree-of-freedom systems such as maximum S_a, average S_a, SRSS S_a, Geomean of S_a, *GMRotI50* S_a, *RotD50* S_a, etc. *RotD50* S_a (Boore 2010) is the current state-of-art intensity measure (IM) that is adopted by all current GMPMs and is expressed as *RotDpp*, where *Rot* indicates the rotation of the two orthogonal components of the ground motion, *D* indicates the period dependency, and *pp* corresponds to the percentile value (mainly limited to the 50th percentiles, i.e. the median value).

The GMPMs mainly express the IM as a function of predictor variables such as Earthquake Moment Magnitude M_w, Closest site-to-source distance R_{rup}, Shear-wave velocity over the upper 30m of the site V_{s30}, etc. Abrahamson *et al.* (2014) proposed the ASK14 ground motion model for shallow crustal earthquakes in active tectonic regions. It included regional differences between California, Japan, China, and Taiwan. Also, Chiou and Youngs (2014) updated their 2008 version of the model and incorporated regional differences in far-source distance and site effects between active tectonic regions in their CY2014 model. Campbell and Bozorgnia (2014) included regionally independent geometric attenuation, regionally dependent anelastic attenuation, and magnitude-dependent aleatory variability in their CB14 model and corrected the bias of their 2008 version.

As the parametric models with functional forms require pre-informed physical interpretation of seismological and geotechnical modeling, researchers oftentimes prefer data-driven non-parametric modeling techniques for developing GMPMs. These models do not require any predefined formulas and make good use of advanced statistical techniques to develop predictive models for the ground motion intensity measures. Tezcan and Cheng (2012) used support vector regression (SVR) which maps data points into a high dimensional feature space to build their non-parametric model to predict response spectra for 13 periods between 0-4s. Recently, Dhanya and Raghukanth (2017) adopted an artificial neural network (ANN) in combination with a Genetic Algorithm (GA) to train a data-driven model to predict PGA, PGV and spectral accelerations at 26 periods between 0.01 and 4s.

However, one of the drawbacks that these GMPM models possess is that they predict the target response at each period independently. In other words, the fixed functional form has different values of coefficients for predicting S_a at different periods of interest, and since the spectrum belongs to the same earthquake event, the S_a predictions for different periods must be correlated. Baker and Jayaram (2010) measured these correlations and concluded that the observed correlations are not sensitive to the choice of GMPMs. They proposed a functional form for a variety of correlation predictions that is valid for periods from 0.01 to 10s. Using this correlational structure a different form of Hazard Spectrum was introduced, known as Conditional Spectrum, for the structural and hazard analysis. Furthermore, Eads *et al.* (2016) developed a metric for quantifying the ground motion intensities called *SaRatio* which is the ratio between S_a at the first period and the average spectral value over a period range. They showed that *SaRatio* is a much better predictor of collapse intensity of a structure than other spectral shape metrics.

This study aims at developing a data-driven non-parametric GMPM model to predict the vector series of S_a (entire Response Spectra) rather than independently predicting the value of S_a at each period separately. Therefore in this study, to incorporate the higher-order spectral dependencies of S_a, the framework of Recurrent Neural Network (RNN) is adopted along with optimation techniques of Evolution Strategy. In particular, under the RNNs, a Long Short Term Memory (LSTM) network is used to estimate a correlated vector of S_a corresponding to 25 periods (periods ranging from 0.1 to 5s) using the input event parameters of Moment Magnitude M_w, Closest site-to-source distance R_{rup}, Shear-wave velocity over the upper 30m of the site V_{s30}, and focal mechanism F which describe the physics of the rupture. Furthermore, the uncertainties are also calibrated in terms of 25x25 covariance matrices that are estimated for both intra-event and inter-event effects by maximizing the log-likelihood calculated from the discrepancy between the spectra estimated by RNN framework and the recorded data.

2. Ground Motion Database

A subset of the PEER NGA-West2 (Timothy *et al.*, 2014) database containing the bi-directional ground motion acceleration records along with the site information, source information, event parameters, and ground motion intensity measures (IMs) is selected as the ground motion database for this study. Some of the records from the database are eliminated based on the exclusion criteria given below:

© The 17th World Conference on Earthquake Engineering

1) Earthquakes that lack any event such as Rake Angle (Focal Mechanism), M_w, etc.
2) Recordings for stations that lack the information about site parameters such as V_{s30}.
3) Earthquakes with a questionable hypocentral depth that are greater than 20 km or less than 1 km.
4) Earthquakes do not belong to shallow active crustal regions.
5) Recordings that are not representative of free field ground motion.
6) Earthquakes with fewer than 5 recordings when M_w<5 and with fewer than 3 recordings when 5< M_w <6.5.
7) Recordings that lack any one horizontal component.
8) Recordings having R_{rup} greater than 80 km.
9) Recordings from aftershocks.

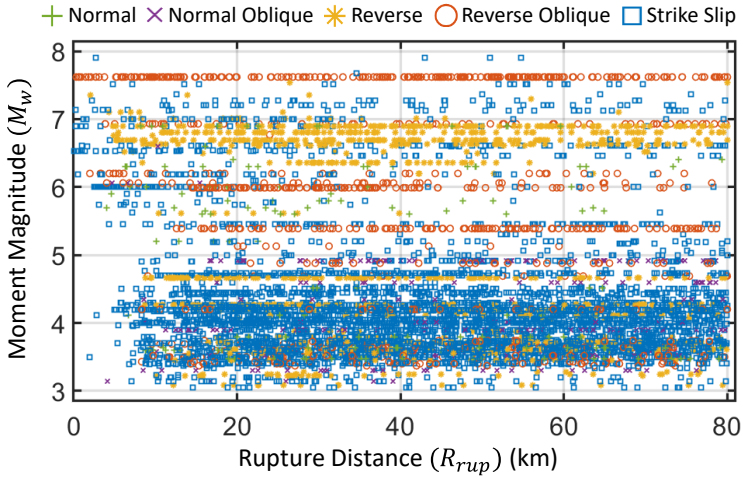


Fig 1 – Magnitude (M_w) and Distance (R_{rup}) details of the selected Ground Motions

Based on the above-mentioned exclusion criteria, 6,947 recordings with 277 earthquakes are finally selected. The earthquakes are classified into 5 fault mechanisms which include Normal (17 earthquakes with 185 recordings), Normal Oblique (12 earthquakes with 245 recordings), Reverse (34 earthquakes with 1,057 recordings), Reverse Oblique (23 earthquakes with 910 recordings), and Strike-Slip (191 earthquakes with 4,550 recordings). The magnitude and distance details of the selected motions are shown in Fig 1. The statistical details of the other parameters obtained from the metadata are listed below:

1) Moment magnitude (M_w) ranges from 3.05 to 7.9.
2) Closest distance to the fault rupture plane (R_{rup}) ranges from 0.07 to 80 km.
3) Joyne-Boore distance to the surface projection of fault rupture plane (R_{jb}) ranges from 0.02 to 80 km.
4) Closest distance to the surface projection of the top of the fault rupture plan (R_x) ranges from -79.93 to 98.57 km.
5) Depth to the top of the fault rupture plane (Z_{TOR}) ranges from 0 to 19.54 km.
6) Average dip angle (δ) of the fault rupture plane ranges from 10 to 90 degrees.
7) Average rake angle (λ) of the fault rupture plane ranges from -180 to 180 degrees.
8) Hypercentral depth measured from sea level (Z_{HYP}) ranges from 0.02 to 20.23 km.
9) Distance to Hypocenter (D_{HYP}) ranges from 2.6 to 160.45 km.
10) Rupture length (L) and Rupture width (W) range between 0.4 to 305 km and 0.3 to 70.4 km, respectively.
11) Shear wave velocity in the top 30m of the profile (V_{s30}) ranges from 116.35 to 2016.13 m/s

3. Parameter Senstivity Analysis

Before developing the RNN framework, it is necessary to identify which source and site parameters possess the highest predictive power to estimate the spectral accelerations. Non-parametric *Random Forests* algorithm

© The 17th World Conference on Earthquake Engineering

is primarily used to conduct the sensitivity analysis of S_a at four periods including 0.2, 0.5, 1.0 and 2.0 secs with respect to all the source and site parameters. The main benefit of using this algorithm as compared to traditional methods of gradient-based sensitivity analysis is that it does not require any closed functional forms to compute the partial derivatives, and due to their non-parametric nature, *Random Forests* has a high power of handling large data sets with higher dimensionality, handling the outliers better and deducing the highly non-linear relationships among the features and target variable. The method of Bagging is used to create randomized decision trees in *Random Forests*, and the bootstrap dataset is created from random sampling (with replacement). Bootstrap datasets that do not contain a particular record from the original dataset are called Out-Of-Bag (OOB) examples (Zhu *et al.*, 2015), and Out-Of-Bag (OOB) estimate for the generalization error is the aggregation of errors of the OOB examples. By first fitting a random forest to the data, OOB error for each data point is recorded and averaged over the forest. The importance of the j^{th} feature is measured by permuting the values of the j^{th} feature among the data and computing the OOB error on this perturbed data set. The importance score for the j^{th} feature is computed by averaging the difference in OOB error before and after the permutation over all trees. Features that produce large values for the normalized (by the standard deviation) score are ranked as more important than features that produce small values. Using the *Random Forests* algorithm, the relative importance (*RI*) of each source and site parameter in predicting the target variables (S_a(T=0.2s), S_a(T=0.5s), S_a(T=1.0s) and S_a(T=2.0s)) is obtained and compared. The results are presented in Fig 2. As can be observed from the subplots for all periods, parameters including M_w, R_{rup}, R_x, Z_{TOR}, and V_{s30} are deemed significant. Specifically, it is observed that the basin effects in terms of Z_{TOR} are observed to become more important for longer periods (1 and 2 secs) while site parameter V_{s30} appears to be important for shorter periods specifically. Source parameters of M_w, R_{rup}, and R_x are consistently observed to be rated significantly by the *Random Forests* algorithm. As the physics of the fault mechanism is widely known to cause differences in the ground motion characteristics, the selected 5 parameters (M_w, R_{rup}, R_x, Z_{TOR}, and V_{s30}) are used along with the fault mechanism to be used as inputs to the RNN framework.

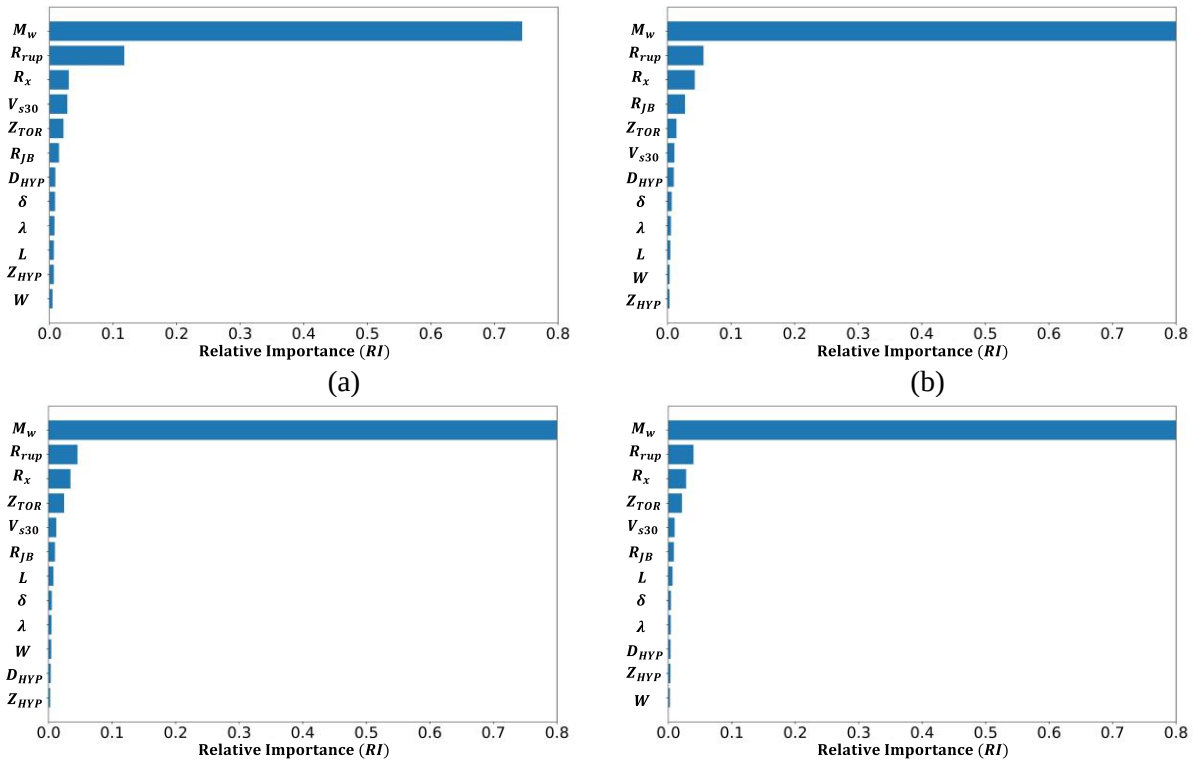


Fig 2 – Relative Importance (*RI*) of source and site parameters to predict: (a) S_a(T=0.2s), (b) S_a(T=0.5s), (c) S_a(T=1.0s), and (d) S_a(T=2.0s)

© The 17th World Conference on Earthquake Engineering

4. Recurrent Neural Networks

Recurrent Neural Network (RNN) is a class of artificial neural network (ANN) that attempts to model time or sequence-dependent behavior. This is performed by feeding back the output of a neural network layer at time t to the input of the same network layer at time $t + 1$. Hence, RNN possesses connections between nodes that form a directed graph along a temporal sequence which allows it to exhibit temporal dynamic behavior. Unlike feedforward neural networks, RNNs can use their internal state (memory) to process input sequences with variable input. RNN not only feeds forward but keeps an internal memory to process the sequences of inputs so that all input vectors are related to each other. Therefore, RNN is the best candidate to train a data-driven model for the prediction of sequential processes, in this case, the S_a Response Spectrum. Since a response spectrum represents the characteristics of the same ground motion, the values of S_a for various periods in the response spectrum are correlated with each other. This can be viewed as a sequence of S_a whose current value depends on the past value. For example, S_a at T=0.3 secs is correlated with the S_a at T =0.2 sec and S_a at T =0.2 secs is correlated with the S_a at T =0.1 sec and so on. The recurrent nature of RNN allows it to perform the same function for each input, copying and sending the data back to the network while producing the output simultaneously.

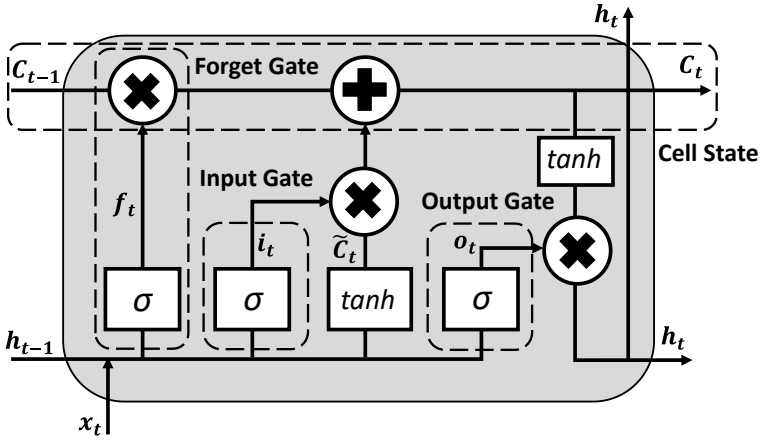


Fig 3- Long Short Term Memory (LSTM) cell structure

Although RNN is capable of tackling dependencies between the steps of the sequences, RNNs are known to have a problem of short-term memory. This means if the sequence is long enough, it will be difficult for them to transfer information from an earlier time step to a later one. During the reverse propagation, RNNs will encounter the problem of vanishing gradient. As the training of neural networks involves the use of gradients to update the weights of nodes, due to the problem of vanishing gradient, the gradient becomes very small over the propagation of time in the sequence and hence the learning does not continue. To tackle this issue, the framework of Long Short Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997) is available. An LSTM is a refined version of traditional RNN which has internal mechanisms called 'gates' that regulate information flow. A typical LSTM structure is a cell state consisting of three gates which are explained below and illustrated in Fig 3.

1) Forget Gate: This determines what information should be discarded or retained from the previous steps in the sequence. Information from the previously hidden state and information from the current input is passed through the sigmoid function (σ). Values between 0 and 1, the closer to 0 means forgetting, the closer to 1 means keeping. The function of the gate is shown in Eq. 1, where the current input x_t and the previous output h_{t-1} are combined using weights W_f and bias b_f with a sigmoid layer.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

© The 17th World Conference on Earthquake Engineering

2) Input Gate: This determines how the cell state (C_t) will be updated, and memory will be modified based on the input. Firstly, the previous hidden state (h_{t-1}) and the current input (x_t) are combined through the sigmoid function (σ) as shown in Eq. 2 using weights W_i and bias b_i to obtain the sigmoid output (i_t). This determines which values are updated by converting the values to be between 0 to 1 with values closer to 0 meaning unimportant and closer to 1 means important. Then using Eq. 3, the hidden state (h_{t-1}) and current input (x_t) are passed to the *tanh* function with weights W_C and bias b_C to take values between – 1 and 1. This helps in regulating the network. Eq. 4 is then used to update the cell state and store the data in terms of the updated cell state (C_t). In Eq. 4, the *tanh* output ($\tilde{C}_t$) is multiplied by the sigmoid output (i_t) and the previous cell state (C_{t-1}) is multiplied by the forget vector (f_t). If f_t is close to 0, C_{t-1} may be discarded in new cell state (C_t). Then the addition of the two products updates the cell state (C_t).

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (4)$$

3) Output Gate: This decides what the next hidden state (h_t) should be and which information will be produced by the state. First, the previous hidden state (h_{t-1}) and current input (x_t) are passed to the sigmoid function with weights (W_o) and bias (b_o) to obtain output (o_t) using Eq. 5. Then the newly modified cell state (C_t) is passed to the *tanh* function and the output is multiplied with the sigmoid output (o_t) using Eq. 6 to determine the information to be carried by the hidden state (h_t). The new hidden state (h_t) is transferred to the next time step.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t \times tanh(C_t) \quad (6)$$

The selected database of the PEER NGA-West2 consisting of 6,947 recordings from 277 events are randomly split into Train and Test sets with 70% of events in the Train set and 30% of events in the Test set. Since the database is heavily dominated by the smaller magnitude events, the Train-Test splitting is conducted in discretized sets of $3 < M_w \leq 4$, $4 < M_w \leq 5$, and so on. This means that 70:30 (Train: Test) random split is done independently for the events in each discretized set and then combined together to form a single Train set and Test set. 10% of the train set is then further used to conduct the neural network training with cross-validation. The neural network framework is trained in epochs using Adam optimizer (Kingsma and Ba, 2014) along with the Early Stopping (Prechelt, 2002) callback to prevent overfitting. The final proposed Neural Network architecture is presented in Fig 4. The source and site inputs to the RNN framework include a vector of 10 values: (1) Magnitude (M_w), (2) Distance to Rupture (R_{rup}), (3) Distance to the surface projection of the top edge of the fault rupture plan (R_x), (4) Depth to the top of the fault rupture plane (Z_{TOR}), (5) Shear wave velocity in the top 30m of the profile (V_{s30}), (6) Fault mechanism (F) represented by one-hot vector for 5 fault mechanisms. Unlike the conventional methods of using discretized values for the discretized classes (as done by Dhanya and Raghukanth, 2017), the 5 classes of fault mechanisms are represented by one-hot vector given in Table 1. Using discretized values for the classes is usually not the correct method to differentiate the classes since the classes that receive arbitrary greater values prioritize the gradient slope in the backpropagation algorithm. Hence to avoid this, the fault mechanisms are differentiated in terms of one-hot vector as given in Table 1. The framework then processes the correlated sequence of S_a using the LSTM layer which estimates the values of S_a while maintaining the spectral correlations. The LSTM layer is then connected to two Dense ANN layers that update the estimations made by LSTM layer so as to obtain better final predictions. The output of RNN framework is the median estimate of 25 points *RotD50* S_a spectrum including S_a values corresponding to periods of 0.1 sec to 1 sec with the interval of 0.1 sec, and 1.2 sec to 3.0 sec with the interval of 0.2 sec, and 3.4 sec to 5 sec with the interval of 0.4 sec.

© The 17th World Conference on Earthquake Engineering

Table 1- One-Hot Vectors for Fault Mechanisms

Mechanism (*F*)	**One-Hot Vector**				
Strike Slip	1	0	0	0	0
Normal	0	1	0	0	0
Reverse	0	0	1	0	0
Reverse Oblique	0	0	0	1	0
Normal Oblique	0	0	0	0	1

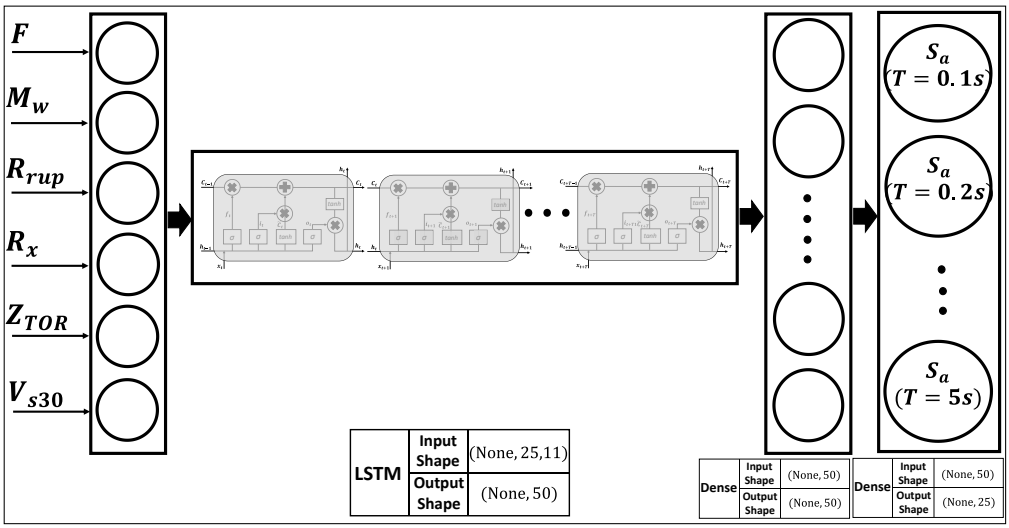


Fig 4 – Proposed Recurrent Neural Network (RNN) Architecture

5. Estimation of Covariance Matrices for Errors

The proposed RNN structure focuses on building a model that estimates a 25x1 vector representing the median values of the response spectrum. Although the nature of RNN effectively addresses the correlation between response spectrum values at multiple periods, the error of the model needs to be dealtcarefully, given that the error terms cannot be independent for the different periods of same spectrum. As the problem statement involves hierarchical data, the error term is partitioned into two parts: inter-event and intra-event effects. η_i, the random effect of the i^{th} event represents inter-event variation, while ε_{ij} stands for the intra-event variation of the j^{th} recording from the i^{th} event. Eq. 7 demonstrates the functional form used to develop the hierarchical model. η_i and ε_{ij} are assumed to be normally distributed with covariances T and Σ, respectively.

$$ln\,Sa_{ij} = RNN(M_i, R_{ij}, \dots) + \eta_i(0, \mathrm{T}) + \varepsilon_{ij}(0, \Sigma)\; ln\,y_{ij} = RNN(M_i, r_{ij}, \dots) + \eta_i + \varepsilon_{ij} \quad (7)$$

T and Σ are the 25×25 covariance matrices that correlate the error terms for the different periods of the spectrum. The diagonal elements of these two matrices represent the inter-event (τ^2) and intra-event variances (σ^2) of the 25 periods, while the off-diagonal element at the i^{th} row and the j^{th} column are $\tau_i\tau_j\rho_{\tau ij}$ and $\sigma_i\sigma_j\rho_{\sigma ij}$ for T and Σ, respectively. $\rho_{\tau ij}$ and $\rho_{\sigma ij}$ are correlations between the i^{th} period and j^{th} period for $\boldsymbol{\tau}^2$ and $\boldsymbol{\sigma}^2$ matrices, respectively. To estimate the optimal covariance matrices T and Σ, the log-likelihood function given in Eq. 8 is maximized using heuristic gradient-free optimization method, namely Covariance Matrix Adaptation Evolution Strategy (CMA-ES) (Hansen, 2010), where y and μ are true and predicted values of S_a, respectively, and C is the covariance matrix. The first term in Eq. 8 is a constant, the second term is the determinant of the covariance matrix, and the third term accounts for the discrepancy between the true response and the predicted values of S_a. Since the data involves heteroscedasticity, following the suggestions of Campbell and Bozorgnia (2014), the inter-event and intra-event covariances matrices T and Σ are estimated separately for $M_w \leq 4.5$ and

© The 17th World Conference on Earthquake Engineering

$M_w \geq 5.5$ and for the values between the two groups, linear interpolation can be performed. Hence, four 25x25 covariance matrices (T and Σ both for $M_w \leq 4.5$ and $M_w \geq 5.5$) are proposed to be used along with the RNN framework for the estimation of the ground motion spectrum using the source, event and site parameters.

$$ln\,L = \frac{N}{2} ln(2\pi) - \frac{1}{2} ln|C| - \frac{1}{2}(y-\mu)^T C^{-1}(y-\mu)\; ln\,L = \frac{N}{2} ln(2\pi) - \frac{1}{2} ln|C| - \frac{1}{2}(y-\mu)^T C^{-1}(y-\mu) \tag{8}$$

CMA-ES is widey used to solve non-linear non-convex optimization problems. In CMA-ES, several particles are randomly generated, and each particle represents a particular setting of parameters that are random variables of the objective function to be minimized or maximized. The mean of the distribution of the particles as well as the covariance matrix of the distribution of the particles are both updated during search steps until the algorithm converges at the global minimum as illustrated in Eq. 9 where m_{t+1} stands for the mean of the distribution at step t+1, and x_i stands for a setting of parameters of the i^{th} particle. The mean of the distribution m_{t+1} at step t+1 is updated based on the mean m_t at step t and the weight w_i, which is the weight of the i^{th} particle proportional to the objective function of the i^{th} particle. μ in Eq. 10 stands for the total number of particles. Eq. 10 illustrates the update of the covariance matrix of the distribution C_{t+1}at step t+1. $c_1$and $c_\mu$are both hyperparamters that affect the rate of convergence. p_{t+1} is the path at step t+1 and $\sigma_t$is a normalization factor at step t , which are updated as shown in Eq. 11 and Eq. 12, respectively. c and μ_w and $c_\sigma$and $d_\sigma$in Eq. 11 and Eq. 12, respectively, are also hyperparameters of the algorithm. $E||N(0,I)||$ in Eq. 12 represents the expectation of the Euclidean norm of the normal distribution $N(0, I)$ ands_{t+1} in Eq. 12 is the step size that is updated based on Eq. 13, where matrix B is formed from the normal distribution $N(0, I)$.

$$m_{t+1} = \sum_{i=1}^{\mu} w_i x_i = m_t + \sum_{i=1}^{\mu} w_i (x_i - m_t) \tag{9}$$

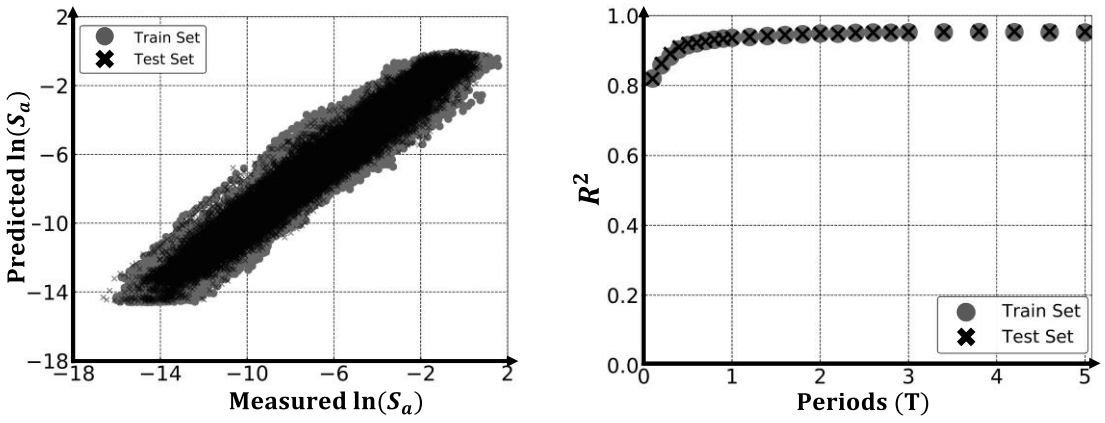
$$C_{t+1} = (1 - c_1 - c_\mu) C_t + c_1 p_{t+1} p_{t+1}^T + c_\mu \sum_{i=1}^{\mu} w_i \left(\frac{x_i - m_t}{\sigma_t}\right) \left(\frac{x_i - m_t}{\sigma_t}\right)^T \tag{10}$$

$$p_{t+1} = (1-c) p_t + \sqrt{c(2-c)}\sqrt{\mu_w}\,\frac{m_{t+1} - m_t}{\sigma_t} \tag{11}$$

$$\sigma_{t+1} = \sigma_t\, exp\left(\frac{c_\sigma}{d_\sigma}\left(\frac{||s_{t+1}||}{E||N(0,I)||} - 1\right)\right) \tag{12}$$

$$s_{t+1} = (1 - c_\sigma) s_t + \sqrt{c_\sigma(2-c_\sigma)}\sqrt{\mu_w} B \sum_{i=1}^{\mu} w_i x_i \tag{13}$$

6. Model Performance



(a) (b)

Fig 5 – (a) Predicted vs Measured S_a for all 25 periods for Train and Test sets, (b) Values of R^2 for the 25 periods for Train and Test sets

© The 17th World Conference on Earthquake Engineering

The proposed RNN framework is evaluated firstly by comparing the actual (measured) values of S_a with the median predictions of S_a made by the RNN framework. This is presented in Fig 5a where the predicted S_a is plotted against the measured values of S_a for both Train and Test sets for all periods combined. As can be observed from the figure, the RNN framework shows good predictions for all ranges of S_a for both Train and Test sets. As mentioned in section 3, the Train and Test sets are split in discretized based on Magnitudes, hence it can be seen from the figure both Train and Test sets contain values of S_a that cover the entire range of S_a. It can be observed from the figure that the scatter between Measured *vs* Predicted S_a for both Train and Test sets tend to follow a 1:1 line which leads to the conclusion that the statistical performance of the RNN framework is satisfactory. Furthermore, the performance of the RNN structure is tested by checking its goodness-of-fit using the statistical measure of R^2. Fig 5b shows the values of R^2 for the 25 selected periods for both Train and Test sets. As can be observed from the figure, the values of R^2 for all periods tend to be greater than 0.8 which shows the high predictive power of the proposed RNN framework. Also, as the R^2 from the Test set is observed to be very close to the R^2 of Train set, it can be concluded that the RNN framework is not overfitted to the database and can be effectively used for future predictions and hazard analysis.

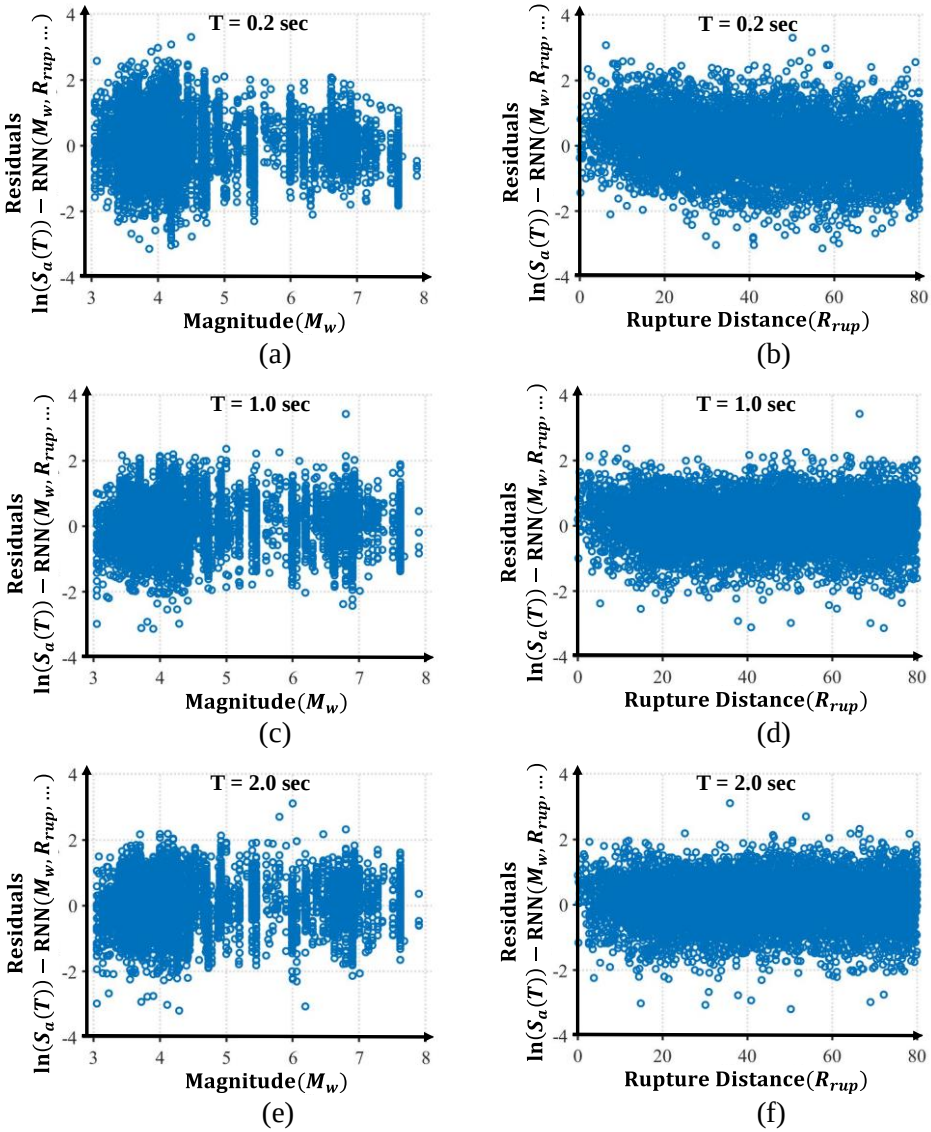


Fig 6 – Residuals of predictions using proposed RNN framework for: (a) $S_a(T=0.2s)$ for M_w, (b) $S_a(T=0.2s)$ for R_{rup}, (c) $S_a(T=1.0s)$ for M_w, (d) $S_a(T=1.0s)$ for R_{rup}, (e) $S_a(T=2.0s)$ for M_w, (f) $S_a(T=2.0s)$ for R_{rup}

© The 17th World Conference on Earthquake Engineering

The performance of the model is further evaluated by checking the residuals obtained from the RNN framework to understand whether the residuals show any observable pattern against Magnitude (M_w) and Rupture Distance (R_{rup}). Fig 6 shows the residuals of the predictions made by the RNN framework against the corresponding Magnitude (M_w) and Rupture Distance (R_{rup}) for periods of 0.2, 0.5, 1.0 and 2.0 secs. Since the residuals are further bifurcated into inter-event (η_i) and intra-event (ε_{ij}), the overall residuals represent the *SRSS* of the inter-event (η_i) and intra-event (ε_{ij}) residuals. As can be observed from Fig 6a to 6f, for all 3 periods (0.2, 1.0, and 2.0 secs), the residuals tend to be normally distributed with mean equal to zero and no trends seem to appear for both Magnitude (M_w) and Rupture Distance (R_{rup}). It can be further observed from Fig 6 that the residuals seem to have smaller sigma for events with Magnitude (M_w) > 5.5 and larger sigma for the events with Magnitude (M_w) < 4.5.

7. Spectral Comparisons

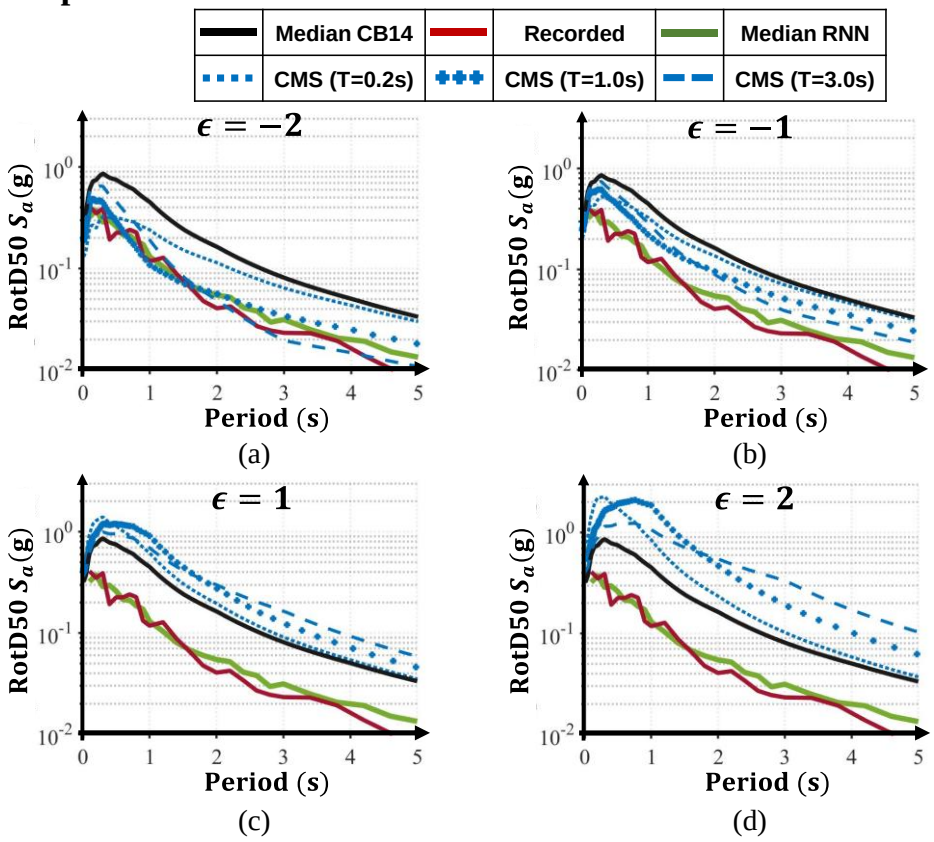


Fig 7 – Spectral Comparsions for ground motion arising from Strike-Slip mechanism with M_w = 6.0, R_{rup} = 9.67 km and V_{s30}= 320.4 m/s for CMS of : (a) $\epsilon = -2$, (b) $\epsilon = -1$, (c) $\epsilon = 1$ and (d) $\epsilon = 2$

In this section, earthquake scenerios are selected from the Test set of the NGAWest2 database and then ground motion spectra generated from the proposed RNN framework are compared against the spectra generated from CB14 (Campbell and Bozorgnia, 2014) GMPE, true recorded spectra, and Conditional Mean Spectra (CMS) (Baker and Jayaram (2010)) developed using CB14, conditioned on three periods ($T = 0.2$ s, $T = 1$ s, and $T = 3$ s) and four epsilons ($\epsilon = -2$, $\epsilon = -1$, $\epsilon = 1$, and $\epsilon = 2$). Fig 7 shows the spectral comparions for ground motion arising from Strike-Slip mechanism with M_w = 6.0, R_{rup} = 9.67 km and V_{s30} = 320.4 m/s. Fig 7a, 7b, 7c and 7d show the recorded spectrum alongwith the spectrum generated from the RNN framework with three CMS conditioned on $\epsilon = -2$, $\epsilon = -1$, $\epsilon = 1$, and $\epsilon = 2$, respectively. As can be observed from the figure, the spectrum generated from the RNN framework lies very close to the true recorded spectrum. Since the CMS are developed conditioned on the period (T) and ϵ, the four sub-figures demonstrate the CMS for on $\epsilon = -2$, $\epsilon = -1$, $\epsilon = 1$, and $\epsilon = 2$ each of which contain three CMS that are conditioned on $T = 0.2$ s,

© The 17th World Conference on Earthquake Engineering

$T = 1$ s, and $T = 3$ s. Based on the Fig 7, it can be observed that having additional constraints on the period (T) and ϵ can cause huge variability in the estimation of the ground motion spectrum. It is observed that CMS conditioned on $\epsilon = -2$ show closer match with respect to the recorded spectrum, however, it can be further observed that the CMS tends to match only at few periods and mainly differ from the recorded spectrum. Similar trends are observed for another scenario (shown in Fig 8) obtained from the Test set for the ground motion arising from Reverse-Oblique mechanism with M_w = 6.9, R_{rup} = 18.33 km and V_{s30}= 663.3 m/s. The spectrum generated from the RNN framework is estimated very close to spectrum of the recorded ground motion and high variability is observed in CMS due to variation in the period (T) and ϵ. In this case, $\epsilon = -1$ generates the three CMS ($T = 0.2$ s, $T = 1$ s, and $T = 3$ s) that are close to the recorded spectrum in the shorter period ranges while tends to o\verestimate for the longer periods. Furthermore, as the GMPEs are developed in terms of average for various earthquake events, it can be observed from both figures 7 and 8, spectra obtained using CB14 highly overestimate the S_a values.

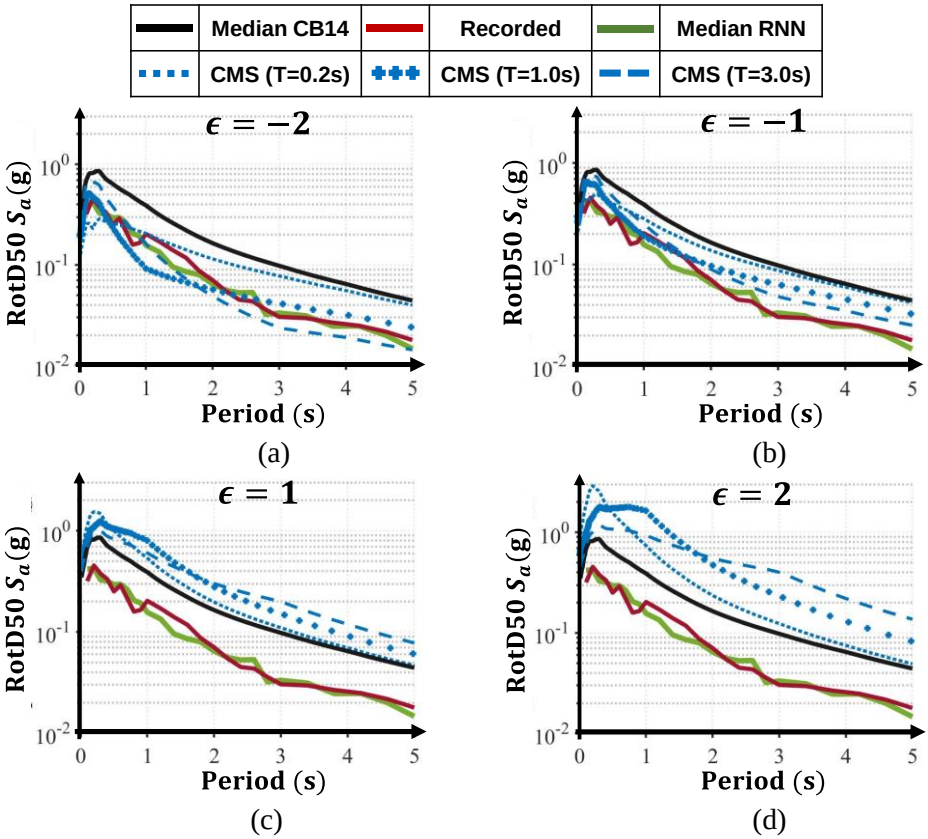


Fig 8 – Spectral Comparsions for ground motion arising from Reverse-Oblique mechanism with M_w = 6.9, R_{rup} = 18.33 km and V_{s30}= 663.3 m/s for CMS of : (a) $\epsilon = -2$, (b) $\epsilon = -1$, (c) $\epsilon = 1$ and (d) $\epsilon = 2$

8. Conclusions

Current seismic hazard and demand analysis are highly based on the utilization of GMPMs. These models are combined with Earthquake Rupture Forecast platforms for probabilistic seismic hazard analysis. Due to the increase in computational resources and outgrowth of machine learning models, the research community has directed these models towards more accurate nonparameteric forms. Though these functional forms are said to be based on the physics of seismic events, the complexity involved in the understanding of the physics of seismic activities has debatably lead to multiple representations of the functional forms. To counteract this, nonparameteric forms can be seen as a good alternative to represent the seismic hazard, especially Neural Networks are argued to have the complex physics embedded in the network and hence can provide a better

© The 17th World Conference on Earthquake Engineering

estimation of the hazard. The hazard is generally represented in the form of an IM, which in the current-state-of-practice is the *RotD50* spectral acceleration of a SDOF system with varying periods. Also, the present GMPMs are not developed explicitly to consider the spectral correlations of IMs for same ground motion.

In this research firstly a sensitivity analysis is conducted through the *Random Forests* algorithm which shows that 5 parameters including M_w, R_{rup}, R_x, Z_{TOR}, V_{s30}, and F are sufficient to accurately define the variability in *RotD50* spectral acceleration of the 7000 mainshock ground motions available in the NGAWest2 database. Using these inputs, an RNN framework is developed to estimate the *RotD50* spectral acceleration for 25 periods while explicitly considering the spectral correlations and dependencies. Furthermore, four 25x25 covariance matrices are developed using heuristic gradient-free optimization methods that estimate the inter-event and intra-event variabilities of the S_a for 25 periods with earthquake magnitudes below 4.5 and above 5.5. Finally, the RNN framework is compared against the CB14 GMPE and CMS by comparing spectra associated with two recorded scenerios. It is observed that the RNN framework show superior spectral match with the recorded ground motions and hence can be easily used for hazard analysis.

9. References

[1] Abrahamson., N.A., Silva, W.J., amd Kamai., R. (2014). Summary of the ASK14 Ground Motion Relation for Active Crustal Regions. *Earthquake Spectra*, 30(3), 1025-1055.

[2] Baker, J.W., and Jayaram, N. (2008). Correlation of Spectral Acceleration Values from NGA Ground Motion Models. *Earthquake Spectra*, 24(1), 299-317.

[3] Boore, D. M., (2010). Orientation-independent, nongeometric-mean measures of seismic intensity from two horizontal components of motion. *Bulletin of the Seismological Society of America* 100 (4), 1830-1835.

[4] Boore, D.M., Stewart, J.P., Seyhan, E., and Atkinson, G.M. (2014). NGA-West2 Equations for predicting PGA,PGV, and 5%-Damped PSA for shallow crustal earthquakes. *Earthquake Spectra*, 30(3), 1057-1085.

[5] Bozorgnia, Y and Campbell, K.W. (2016). Vertical ground motion model fpr PGA, PGV, and linear response spectrum using the UGA-West2 database. *Earthquake Spectra*, 32(2), 979-1004.

[6] Campbell K.W., and Bozorgnia, Y. (2014). Campbell-Bozorgnia NGA-West2 horizontal ground motion model for active tectonic domains. *Earthquake Spectra*, 30(3), 1087-1115.

[7] Chiou, B.S., and Youngs, R.R. (2014). Update of the Chious and Youngs NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra. *Earthquake Spectra*, 30(3), 1117-1153.

[8] Dhanya, J, and Raghukanth, S.T.G. (2017). Ground Motion Prediction Model Using Artificial Neural Network. *Pure and Applied Geophysics*, 175, 1035-1064.

[9] Douglas, J. (2019). Ground motion prediction equations 1964–2019. Accessed 19th Dec 2019. http://www.GMPM.org.uk/GMPMreport2014.pdf

[10] Eads, L., Miranda, E., and Lignos, D. (2016). Spectral shape metrics and structural collapse potential. *Earthquake Engineering and Structural Dynamics*, 45, 1643-1659.

[11] Hansen, N. (2010) Variable Metrics in Evolutionary Computation. *Habilitation a diriger des recherches*, University Paris-Sud.

[12] Hochreiter, S., and Schmidhuber, J. (1996). LSTM can solve hard long time lag problems. *Proceedings of the 9th International Conference on Neural Information Precossing Systems*, December 1996, 473-479.

[13] Jarayam, N., and Baker, J.W. (2010). Correlation of response spectral values in Japanese ground motions. *Earthquake and Structures*, 2(4), 357-376.

[14] Kingma, D.P., and Ba, J. (2014). Adam: A Method for Stochastic Optimization. *3rd International Conference for Learning Representations*. San Diego, 2015.

[15] Prechelt, L. (2002). Early Stopping – But When? *Neural Networks: Tricks of the Trade*. University Karlsruhe.

[16] Tezcan, J., and Cheng, Q. (2012). Support vector regression for estimating earthquake resoonse spectra. *Bulletin of Earthquake Engineering*. 10(4), 1205-1219.

[17] Timothy, D.A., Robert, B.D., and Jonathan, P.S. (2014). NGA-West2 Database. *Earthquake Spectra*, 30(3), 989-1005.

[18] UCERF2, The Uniform California Earthquake Rupture Forecase, Version 2. (2008). U.S. Geological Survey, California Geological Survey, and the Southern California Earthquake Center.

[19] UCERF3, The Uniform California Earthquake Rupture Forecase, Version 3. (2015). U.S. Geological Survey, California Geological Survey, and the Southern California Earthquake Center.

[20] Zhu R, Zeng D, Kosorok MR (2015). Reinforcement Learning Trees. *Journal of the American Statistical Association*. 110 (512): 1770–1784.

© The 17th World Conference on Earthquake Engineering