

## ON THE APPLICATION OF NEURAL NETWORKS TO GROUND MOTION INTENSITY AND CORRELATION MODELLING

G. J. O'Reilly<sup>1</sup>, S. Aristeidou<sup>2</sup> & D. Shahnazaryan<sup>2</sup>

<sup>1</sup> Centre for Training and Research on Reduction of Seismic Risk (ROSE Centre), Scuola Universitaria Superiore IUSS Pavia, Italy, [gerard.oreilly@iusspavia.it](mailto:gerard.oreilly@iusspavia.it)

<sup>2</sup> Centre for Training and Research on Reduction of Seismic Risk (ROSE Centre), Scuola Universitaria Superiore IUSS Pavia, Italy

**Abstract:** *This paper presents the application of neural networks to ground motion intensity modelling. We focus on the development of a generalised ground motion model (GGMM) incorporating several seismic intensity measures (IMs) and quantifying the correlations between them. These range from classical IMs such as spectral acceleration to more novel and advanced IMs recently shown to be much better descriptors of structural performance and seismic risk. A mixed-effects regression approach is adopted to capture the inter- and intra- event variability of the GGMM prediction. Artificial neural network (ANN) is used to perform the regression, which differs from the approach used in many past works and present in many existing ground motion models (GMMs). The correlations between the IMs were also quantified, which allows for a more refined prediction of seismic shaking and a unified treatment of prediction and correlation. This will allow more advanced record selection for non-linear dynamic analyses to be performed and considers several facets of ground shaking currently overlooked in many works. We evaluate the performance of the developed GGMM using several metrics and comparing it to various GMMs currently used and developed with either the classical method or with machine learning methods. The results show that the proposed GGMM exhibits improved performance, while allowing a more consistent cross-correlation between the different IMs. This paper outlines the general methodology, a general overview and discusses possible implications.*

### 1. Introduction

Ground motion models (GMMs) are an essential part of a seismic hazard analysis, regional seismic analysis, structural loss estimation, shake maps, and more, which are included in both the fields of earthquake engineering and seismology. GMMs estimate the mean ground motion intensity and dispersion, given a set of causal rupture parameters (e.g., magnitude, source-to-site distance, etc.). Many different intensity measures (IMs) can be used to characterise the ground motion shaking intensity at the site of interest. There is a growing interest in using cumulative intensity-based IMs (i.e.,  $DS_{595}$ ) together with peak response amplitude-based IMs (i.e.,  $S_a$ ,  $FIV3$ ), which sparked the development of a plethora of GMMs to estimate each type of IM. However, these GMMs (e.g., Bradley, 2011; Afshari and Stewart, 2016; Campbell and Bozorgnia, 2019) predict the IMs independently, with each available GMM being based on a different ground motion database (or at least applying different filtering criteria) and different regression models for the fit. This leads to some degree of heterogeneity, which can be mitigated by developing a generalised ground motion model (GGMM) to estimate all of the different type of IMs collectively (Fayaz et al., 2020). Using independent GMMs to estimate assorted ground motion IMs, like  $S_a$ ,  $S_{avg}$ ,  $FIV3$ ,  $DS_{595}$ , etc., for the same earthquake scenario can introduce unwanted bias and increased variability, which is then propagated into the results of the seismic analysis and risk assessment.

Traditionally, GMMs are developed using parametric functional forms and fitting a set of coefficients based on empirical data. In recent years, researchers have been exploring the potential of data-driven non-parametric

regression techniques for developing GMMs, as for example the work Dhanya and Raghukanth (2018), which has the advantage of not requiring to set up mathematical formulations (i.e., functional forms) that describe best the physical phenomenon. It is acknowledged, however, that relying purely on data-driven approaches is not a perfect solution as these models work well where data is available. In ground motion modelling, we are typically interested in strong shaking which requires data from large magnitude earthquakes, which are less frequent. Hence, there is a danger that utilising these data-driven GMMs to predict intensities that are not well recorded and fitted to, or even beyond the fitting range, that inaccuracies may arise. It is here that robust verification is needed and, in some cases, the physical meaning of the parametric functional forms may be advantageous.

Therefore, in this study the GGMM was developed using an artificial neural network (ANN) framework that can be used to estimate a set of intensity measures (IMs) utilising seismic source, distance, and site parameters as input, which are listed in Section 3. The results and predictions of the proposed GGMM were then presented and compared against the aforementioned recent and well-established GMMs. The following GMMs from the literature were used for the comparisons: Campbell and Bozorgnia (2008) for peak ground displacement (*PGD*), denoted as CB08; Campbell and Bozorgnia (2014) for peak ground acceleration (*PGA*), peak ground velocity (*PGV*), and spectral acceleration,  $S_a$ , denoted as CB14; Dávalos et al. (2020) for filtered incremental velocity, *FIV3*, denoted as DHM20; Dávalos and Miranda (2018) for  $S_{a_{avg}}$ , denoted as DM18; Afshari and Stewart (2016) for  $D_{S_{95}}$ , denoted as AS16. The following section describe the ground motion database utilised, the predictor and response features used in the ANN model followed by an evaluation of the model's performance.

## 2. Strong motion database and filtering

To utilise ANN to fit a GGMM, a dataset of ground motion recordings was first required. The PEER NGA-West2 (Ancheta et al., 2013) database containing bi-directional ground motion acceleration records including site details, source information, and ground motion IMs was adopted. The database was filtered to eliminate some of the records that may be deemed unsuitable for general use based on the criteria given below:

- Only ground motion records from earthquakes with moment magnitude,  $M_w$ ,  $\geq 4.5$  were utilised. Earthquakes of lower magnitude were omitted due to insufficient intensity to induce significant non-linear deformations or structural collapse in engineered buildings without appropriate scaling.
- Recordings from instruments located on free field or below surface or in the first storey of low-rise structures (less than four storeys) were utilised. This was based on the Geomatrix 1<sup>st</sup> letter code of the NGA-West2 flat-file.
- Events with hypocentral depth greater than 20 km were discarded.
- Events recorded on bedrock were discarded based on the upper threshold of 1200 m/s of mean shear wave velocities in the upper 30 meters,  $V_{s,30}$ .
- Recordings with a rupture distance,  $R_{rup}$ , greater than 300 km were discarded.
- Only recordings from strike-slip, reverse, and reverse-oblique events from active shallow crustal tectonic environments were included.
- Only records whose minimum usable frequency of both components was smaller than 0.25 Hz were considered.
- Earthquakes with fewer than 3 recordings were discarded, as they could be considered poorly recorded, and therefore adversely impact the intra-event variability of the model.
- Recordings were considered only if both horizontal components were available.
- Recordings from aftershocks were excluded since most seismic hazard analyses use GMMs based on mainshocks. In this study, a recording is classified as a aftershock if it is defined as 'Class 2' event with  $CR_{JB} < 10$  km according to the criteria given in Wooddell and Abrahamson (2014).

Based on the filtering criteria outlined above, the final ground motion database included 4134 recordings with 96 earthquakes. The earthquakes were classified into five styles of faulting (SOF) including strike-slip (53 earthquakes and 1815 recordings), normal (10 earthquakes and 75 recordings), normal oblique (4 earthquakes and 249 recordings), reverse (20 earthquakes and 1075 recordings) and reverse oblique (9 earthquakes and 917 recordings). Figure 1 displays the  $M_w$ ,  $R_{rup}$  and  $V_{s,30}$  distributions of the filtered database. Additionally, the depth to shear velocity of 2.5 km/s,  $Z_{2.5}$ , when missing for some ground motion recordings, was estimated following the prediction equations of Kaklamanos et al. (2011).

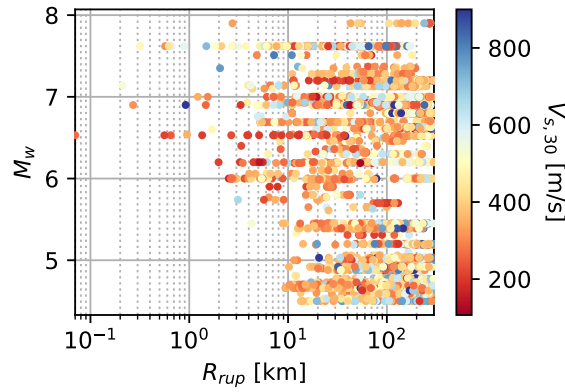


Figure 1.  $M_w$ ,  $R_{rup}$  and  $V_{s,30}$  distribution of the filtered database

### 3. Predictor and response features

Prior to the development and training of Artificial Neural Network (ANN) models, it is essential to identify the predictor and response features. The informed selection of predictor features is essential to the robustness and accuracy of subsequent modelling process. Within the scope of GMMs, past research (e.g., Fayaz et al., 2020) has highlighted the substantial predictive power of moment magnitude,  $M_w$ , and rupture distance,  $R_{rup}$ . In addition, several other source parameters were included for the training of the ANN models. The full list of predictor and response features within this study are given in Table 1. Users are required to provide a value for each of the predictor features shown in the first column and can obtain predictions for any of the response features listed in the second column, essentially making it a generalised GMM for the variety of IMs that can be predicted.

Table 1. List of eventual predictor and response features, along with their horizontal component definition.

Predictor features	Response features
Moment magnitude, $M_w$	$PGA$
Rupture distance, $R_{rup}$ [km]	$PGV$
Hypocentral distance, $R_{hyp}$ [km]	$PGD$
Time-averaged shear-wave velocity to 30m depth, $V_{s,30}$ [m/s]	$D_{5-95}$
Style of faulting, $SOF$	$Sa(T)$
Depth to the 2.5 km/s shear-wave velocity horizon (a.k.a., basin or sediment depth), $Z_{2.5}$ [m]	$FIV3(T)$
Depth to top of fault rupture, $Z_{tor}$ [m]	$Sa_{avg}(T)$
Joyner-Boore distance, $R_{jb}$ [km]	
Distance measured perpendicular to the fault strike from the surface projection of the up-dip edge of the fault plane, $R_x$ [km]	

Response features, or IMs essentially, included the  $PGA$ ,  $PGV$ ,  $PGD$ ,  $D_{5-95}$ , 22 definitions of  $Sa$  at periods ranging from 0.01s to 5.0s, 14 definitions of average spectral acceleration,  $Sa_{avg}$ , at periods ranging from 0.1s to 4.0s given by Equation (1) (Eads et al., 2015), 14 definitions of filtered incremental velocity,  $FIV3$ , at periods ranging from 0.1s to 4.0s given by Equation (1) (Dávalos and Miranda, 2019).

$$Sa_{avg}(c_1T, \dots, c_NT) = \left( \prod_{i=1}^N Sa(c_iT) \right)^{\frac{1}{N}} \quad (1)$$

where  $S_a$  corresponds to the 5%-damped pseudo-acceleration spectral value,  $c_i$  is factor ranging uniformly from 0.2 to 2.0  $N=10$  times. Previous research has shown that this spacing scheme is more efficient than a logarithmic one and that the difference between using 10 or 100 periods is negligible, on average (Eads and Miranda, 2013).

$$FIV3 = \max\{V_{s,max1} + V_{s,max2} + V_{s,max3}, |V_{s,min1} + V_{s,min2} + V_{s,min3}|\} \quad (2)$$

$$V_s(t) = \int_t^{t+\alpha T} \ddot{u}_{gf}(t) dt, \quad \forall t < t_{end} - \alpha T \quad (3)$$

where  $V_s(t)$  is a series of incremental velocities (IVs) estimated using time segments of  $\alpha \cdot T$ ,  $V_{s,max1}$ ,  $V_{s,max2}$ ,  $V_{s,max3}$  are the first, second, and third local largest IVs in  $V_s(t)$ , respectively, and  $V_{s,min1}$ ,  $V_{s,min2}$ ,  $V_{s,min3}$  are the first, second, and third local minimum IVs in  $V_s(t)$ , respectively,  $T$  is the period of interest,  $t_{end}$  is the last instant of time of acceleration time series, and  $\ddot{u}_{gf}$  is the filtered acceleration time series using a second-order Butterworth low-pass filter with a cut-off frequency,  $f_c$ , equal to  $\beta \cdot f$ , where  $\beta$  is a scalar controlling the  $f_c/f$  ratio and  $f$  is  $1/T$ . The parameters  $\alpha$  of 0.7 and  $\beta$  equal to  $T$  are based on the findings of Dávalos et al. (2020).

Ground motions are usually recorded in three orthogonal directions in space, so there is a need to combine these recorded directions into an IM with a specified horizontal component definition. Several horizontal component definitions have been used in the literature to quantify the intensity of a ground motion on single-degree-of-freedom systems based on the two orthogonal horizontal components, such as maximum of the two, average, square-root-of-sum-of squares, geometric mean, *GMRotI50*, *RotD50*. Most modern GMMs use the *RotD50* definition (Boore, 2010), as it is commonly accepted to be the state-of-the-art horizontal component definition for spectral IMs. The *RotD50* definition was adopted here for  $S_a$ , whereas for more advanced IMs (i.e., *FIV3* and  $S_{avg}$ ) and for other IMs (i.e., *PGA*, *PGV*, *PGD* and  $D_{S595}$ ) the geometric mean definition was adopted.

## 4. Model architecture

### 4.1. Fixed-effects with artificial neural network

A “feed-forward” ANN was employed here for prediction of the IMs outlined in Section 3. ANN is a subset of deep learning and is composed of artificial neurons interconnecting an input layer, one or more hidden layers, and an output layer (McCulloch and Pitts, 1943). Each neuron in an ANN performs a simple computation, where it receives a signal, applies an activation function, and passes the result through the hidden layers to the output layer, hence the term “feed-forward”. While the neurons in the hidden layers process the information, the neurons in the input layer simply transmit the input data, and the neurons in the output layer provide the final outputs, or within the scope of this study, the predictions of IMs of interest. Each connection has an associated synaptic weight representing the strength of the connection. Similarly, neurons of the network are associated with a bias term, which determines the threshold at which a neuron is activated. The synaptic weights of the connections along with biases of the neurons represent the parameters of the neural network which are adjusted during the training process to optimise the performance of the network. The synaptic weights are used as the multipliers of the outputs of the previous layer and the bias is a constant added to the outputs before passing through the activation function. The training is typically done through a technique called backpropagation, which uses a gradient descent optimisation (Kiefer and Wolfowitz, 1952), where the network tries to minimise the difference between its predictions and the actual target values in the training dataset. For detailed description of neural networks, readers are referred to Haykin (2009). The schematics of the chosen ANN architecture is shown in Figure 2, and the general expression to predict each IM is as per Equation (4).

$$\log_{10}(IM_r) = f_{linear} \left[ b_r + \sum_{h=1}^{40} W_{h,r} \cdot f_{softmax} \left( b_h + \sum_{p=1}^9 W_{p,h} X_p \right) \right] \quad (4)$$

where  $X_p$  is the predictor feature  $p$  (Table 1),  $W_{p,h}$  is the weight of the connection between predictor neuron  $p$  and hidden neuron  $h$  for the predictor  $X_p$ ,  $b_h$  is the bias of the hidden neuron  $h$ ,  $W_{h,r}$  is the weight of the connection between hidden neuron  $h$  and response neuron  $r$ ,  $b_r$  is the bias of the response neuron  $r$  (from 1 to 54),  $f_{softmax}$  and  $f_{linear}$  are the activation functions of the hidden and response layers, respectively.

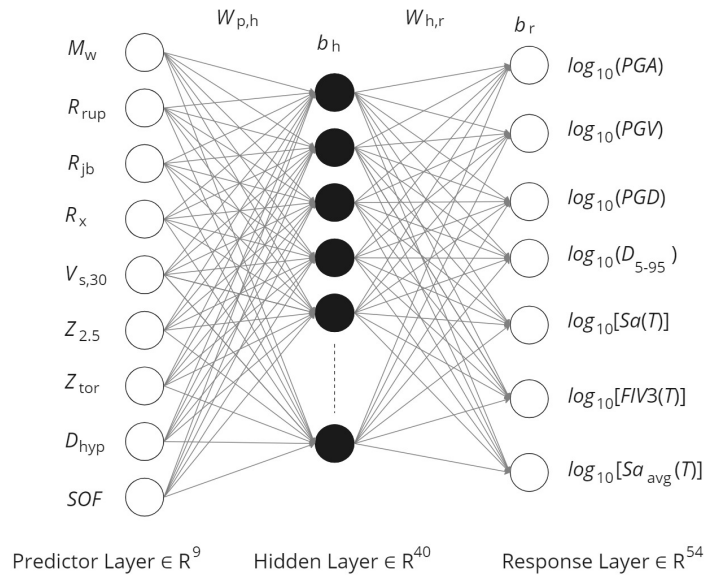


Figure 2. Architecture of the ANN ground motion model

The ANN training was performed within a Python environment, using the open-source TensorFlow library (Abadi et al., 2015). The step-by-step implementation, including dataset processing, training of ANNs and generation of predictions, are outlined herein. Dataset feature processing and selection of ANN parameters and functions is described as follows:

- Predictor feature engineering: *MinMax* normalisation is adopted to ensure that predictor features are on a similar scale, hence have a comparable influence on the model's learning process (e.g., magnitude tends to range between  $M_w = 4.5 - 8$ , but soil properties can vary between  $V_{s,30} = 200 - 1200$  m/s). This can improve the convergence of the training process and make it less sensitive to the scale of predictor features. Additionally, the initialisation of weights can be more effective, which will facilitate faster convergence and prevent gradients vanishing or exploding issues. The scaling was done using a range of 0.1 to 0.9 instead of 0 and 1 to maintain a small buffer at both ends of the range, where the data may contain values close to the minimum and the maximum but is not viewed as an outlier. For what regards *SOF*, one-hot encoding was applied.
- Response (IMs) feature engineering: Similar to predictor features, the response features are scaled to span similar ranges. Typically, a  $\log_{10}$  transformation is applied to the vector of IMs. The  $\log_{10}$  transformation limits the response parameter range more than a natural logarithm transformation, and therefore a more robust fit could be achieved (i.e., better performance metrics).
- Number of hidden layers and neurons: A single hidden layer was employed following a trial-and-error approach, which demonstrated that using just one hidden layer was adequate for making predictions. The input layer consisted of 9 neurons, matching the number of predictor features, while the output layer consisted of 54 neurons, corresponding to the number of considered IMs. Concerning the number of neurons in the hidden layer, 40 neurons were chosen, as it produced the model's optimal predictive performance (considering the chosen performance metrics and eventual model dispersion); using fewer or more neurons led to either underfitting or overfitting, respectively.
- Activation functions: Calculates the output of a neuron. Given the nature of regression problems, normalised exponential (also known as *softmax*) and linear activation functions were considered in the hidden and output layer, respectively, based on hyperparameter tuning described later. *Softmax* proved to work better with the range of predictor features, given the *MinMax* normalisation. Furthermore, *softmax* in the hidden layer, introduces non-linearity, which enables the network to learn complex patterns from the data.
- Optimisation algorithm and loss function: The loss function employed for optimisation was the mean squared error (MSE) given by Equation (5), and its minimisation was accomplished through the use of the adaptive moment (ADAM) estimation algorithm (Kingma and Ba, 2014), a selection made following hyperparameter tuning. Additionally, the coefficient of determination,  $R^2$ , given by Equation (6), is used

to determine how well the variation of response features is explained by predictor features in a regression model.

- Training and testing sets: Prior to model training, the filtered dataset from Section 2 was randomly split into training and testing sets using an 80-20 ratio.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (5)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (6)$$

where  $y_i$  is the  $i^{\text{th}}$  observed value,  $\hat{y}_i$  is the  $i^{\text{th}}$  predicted value, and  $\bar{y}$  is the mean value of  $n$  data points.

The next step of the implementation involves training the ANN. To assess the model's performance, a fivefold cross validation (Picard and Cook, 1984) was employed. The training set was randomly partitioned into five equal-sized sets. Five separate ANNs were trained, each using four of the subsets for training and the remaining fifth subset for model prediction validation. The procedure ensures, that each subset takes on the role of the validation set for its respective training. The fixed-effect regression metrics for cross-validation were computed as the average of the results from five ANNs. Furthermore, Bayesian optimisation (Moćkus, 1975) was employed to determine the optimal hyperparameters for the ANN regression model. The objective within the context of this study was to minimise the MSE of the fixed-effects regression by exploring a range of hyperparameters. A summary of the hyperparameters considered are provided below:

- Batch size from 8 to 128: helps balance computational efficiency and model performance. With smaller batch sizes better model generalisation can be achieved, however it can be computationally insufficient, as more updates are needed to process the entire dataset. In contrast, larger batch sizes aid in accelerating the training, but can hinder model generalisation, hence, the model is more prone to overfitting.
- Training epochs from 50 to 200: during each epoch, the model passes through all training samples and updates its parameters (weights and biases) based on the loss incurred when making predictions. The updates try to minimise the error and improve the model's performance. While the optimal number of training epochs can improve the model's ability to generalise, with increasing number of epochs, overfitting may incur. Therefore, early stopping was implemented as a preventive measure against overfitting, which automatically halts training if the model stops improving over a span of 20 consecutive epochs.
- Optimisation algorithm: The following optimisation algorithms were considered: ADAM; root mean square propagation (RMSprop); stochastic gradient descent (SGD); adaptive gradient descent (Adagrad); adaptive learning rate (Adadelta); a variation of ADAM (Adamax); combination of Nesterov accelerated gradient and Adam (Nadam); follow the regularised leader (Ftrl).
- Learning rate of the optimisation algorithm from  $0.5 \times 10^{-3}$  to 0.05: controls the step size during weight updates and influences the convergence speed and stability of the model.
- Activation function of hidden layer: The following activation functions were considered: linear; rectified linear unit (ReLU); leaky ReLU; exponential linear unit (ELU); scaled ELU; *softmax*; hyperbolic tangent (tanh).

The approach was utilised to provide a comprehensive evaluation of the model's performance while mitigating the potential risks associated with overfitting (high variance) and underfitting (high bias). The hyperparameters that yielded the best model performance are as follows: *softmax* and linear activation functions for the hidden and response layers, respectively; learning rate of  $0.98 \times 10^{-3}$ ; a batch size of 32; and 100 training epochs.

#### 4.2. Mixed-effects regression

The general form of the generalised ground motion model is given as:

$$\log_{10} IM_i = f_i(\mathbf{X}, \boldsymbol{\theta}) + \delta b_i \tau_i + \delta w_i \varphi_i \quad (7)$$

where  $\log_{10}(IM_i)$  is logarithm with base 10 of the  $i^{\text{th}}$  IM;  $f_i(\mathbf{X}, \boldsymbol{\theta}) = \mu_{\log_{10} IM_i | \mathbf{X}, \boldsymbol{\theta}}$  is the predicted mean output from the ANN model, taking as input a set of GM causal features (e.g.,  $M_w$ ,  $R_{rup}$ , etc.), denoted as  $\mathbf{X}$ ;  $\boldsymbol{\theta}$  are the 'calibrated coefficients' of the ANN model (i.e., synaptic weights and biases);  $\delta b_i$  and  $\delta w_i$  are the normalised

inter- and intra-event residuals of  $IM_i$ , respectively;  $\tau$  and  $\varphi$  are the inter- and intra-event logarithmic standard deviations. The main metric to evaluate the performance of the model is the total standard deviation,  $\sigma$ . In the dispersion model of this GMM,  $\sigma$  was kept magnitude-independent since recent GMMs noticed only minor dependencies on  $M_w$  and only for  $M_w < 5.5$  (Boore et al., 2014; Campbell and Bozorgnia, 2014). To calculate  $\sigma$ , must first segregate the total residuals between inter- and intra-event residuals, which can be treated as normal variables that ideally should follow a normal distribution with zero mean and standard deviations  $\tau$  and  $\varphi$ , respectively (Atik et al., 2010). If the inter- and intra-event residual are assumed to be mutually independent, then the total standard deviation can be calculated as the sum of their variances, given in Equation (8).

$$\sigma = \sqrt{\tau^2 + \varphi^2} \quad (8)$$

Taking advantage of this assumption and the better understanding of these two different sources of uncertainty, Abrahamson and Youngs (1992) proposed an one-step mixed-effect regression algorithm, using the maximum likelihood approach, to compute the variances  $\tau^2$  and  $\varphi^2$ . This algorithm is an iterative procedure in which mixed-effects, variances, and model parameters are computed successively. This procedure is now widely applied for the development of a GMM, and so herein it is adapted to the ANN model. The adapted algorithm is based on the procedure proposed in Abrahamson and Youngs (1992) and similar to the one used in Derras et al. (2014) and can be summarised as follows:

1. Estimate the initial set of ANN model parameters (i.e.,  $[W_{1, 2, \dots, i}]$  and  $\{b_{1, 2, \dots, i}\}$ ) using fixed-effect training procedure.
2. Estimate  $\tau^2$  and  $\varphi^2$  from  $[W_{1, 2, \dots, i}]$  and  $\{b_{1, 2, \dots, i}\}$ , by maximising the log-likelihood function given in equation (7) of Abrahamson and Youngs (1992).
3. Given  $[W_{1, 2, \dots, i}]$ ,  $\{b_{1, 2, \dots, i}\}$ ,  $\tau^2$  and  $\varphi^2$ , estimate the random inter-event term  $\eta_i$  given in equation (10) of Abrahamson and Youngs (1992).
4. Estimate the new  $[W_{1, 2, \dots, i}]$  and  $\{b_{1, 2, \dots, i}\}$ , using fixed-effects training procedure for  $(\log_{10}(Y) - \eta_i)$ .
5. Repeat steps 2, 3, and 4 until the termination criterion is satisfied. The adopted termination criterion is 0.1% in terms of the difference between two successive likelihood values.

## 5. Model performance

### 5.1. Performance metrics and comparison with other GMMs

The performance of the ANN model can be evaluated by comparing the empirical (i.e., recorded) values of IMs with their corresponding median predictions of the model using a variety of metrics, which in this case were the MSE and  $R^2$ . The resulting average MSE for the model, determined through a fivefold cross-validation with the optimal parameters, obtained as described in Section 4.1, were found to be 0.093 for the training set and 0.095 for the validation set. Finally, the model corresponding to the optimal parameters, after passing through the mixed-effects regression, was evaluated using the 20% unseen testing set and the regression metrics: MSE and  $R^2$ , associated with each IM are reported in Figure 3. The eventual average MSE of all IMs were 0.093 for the training set and 0.095 for the test set. It is noteworthy that the average training and validation set MSE after cross-validation with the optimal parameters gives the same average training and validation set MSE, respectively, further validating the model's accuracy. From Figure 3, it can be seen that the  $R^2$  of the testing set is, in general, only slightly lower than that of the training set. At the same time, both values are high, indicating that the model has high predictive power and avoids overfitting. The same applies to the MSE metric, which is inversely proportional to the model's predictability.

To evaluate the GGMM's performance visually with respect to the available data and also some comparable GMMs available in the literature, some visual comparisons were plotted. Figure 4 shows the magnitude amplification of  $Sa(2.0s)$  for two different rupture distance bins compared with the GMM of Campbell and Bozorgnia (2014) labelled as CB14. It can be seen that the predicted values are generally close to the cloud mean, with a minor deviation in medium-range distances (i.e.,  $50 \text{ km} \leq R_{rup} \leq 100 \text{ km}$ ) in high magnitudes, and the CB14 values also illustrate this deviation trend but are generally close to the values predicted from the proposed model, giving confidence to the proposed GGMM. Likewise, Figure 5 presents the distance attenuation of  $FIV3(1.0s)$  for two different magnitude bins and it is compared with the model of Dávalos et al. (2020), which to date is the only other GMM available for this IM. It can be seen that the proposed model does well in capturing the trends of the cloud median, while the DHM20 model generally predicts higher values, especially in lower magnitudes. The same results are illustrated in Figure 6, but for  $DS_{595}$  and compared with

Afshari and Stewart (2016). While the trends between the two models are the same, the AS16 model predicts somewhat lower values of significant duration than the proposed model. This difference is speculated to be because of the different database filtering criteria to exclude recordings with unreasonably large durations polluted by high-frequency noise, even though the same database was used (i.e., NGA-West2).

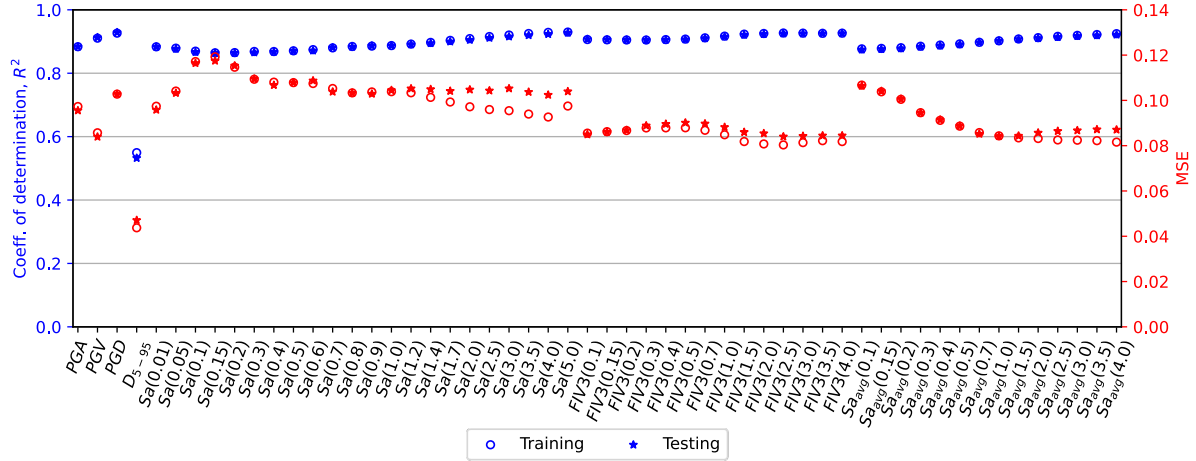


Figure 3. Training and testing MSE and  $R^2$  values of the ANN model post fixed-effects

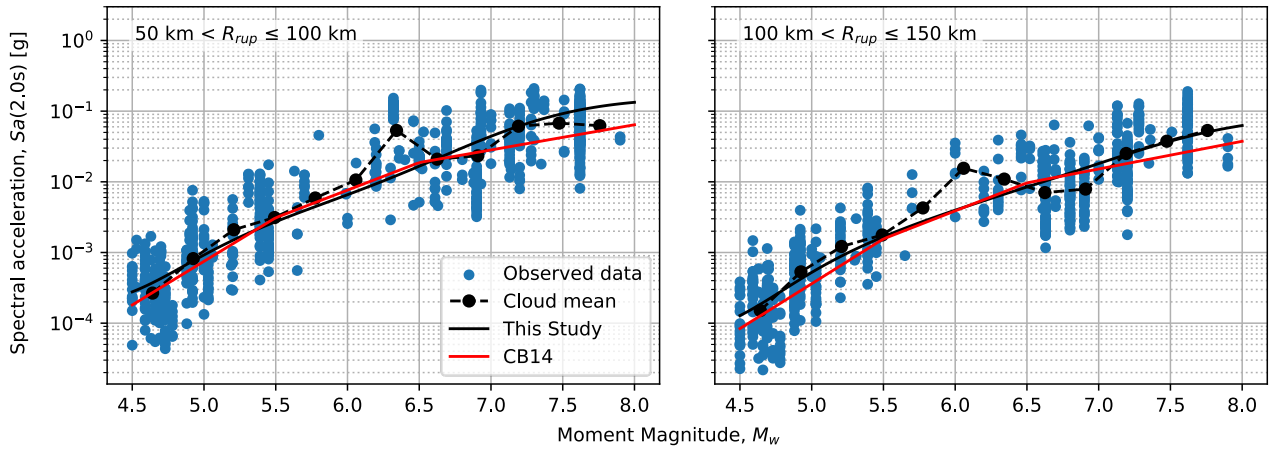


Figure 4. Magnitude amplification plots of  $Sa(2.0s)$  for two different distance bins

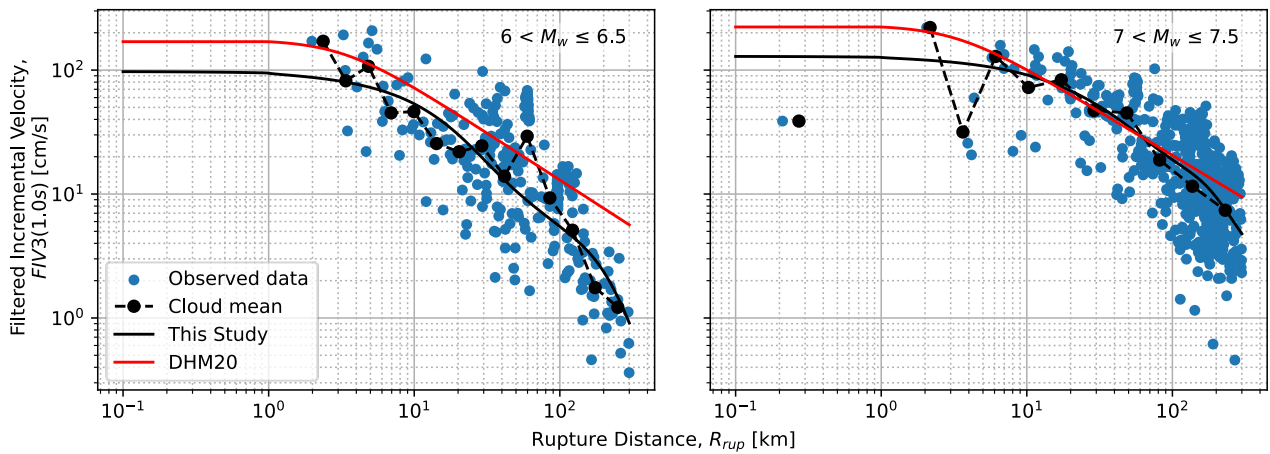


Figure 5. Distance attenuation plots of  $FIV3(1.0s)$  for two different magnitude bins



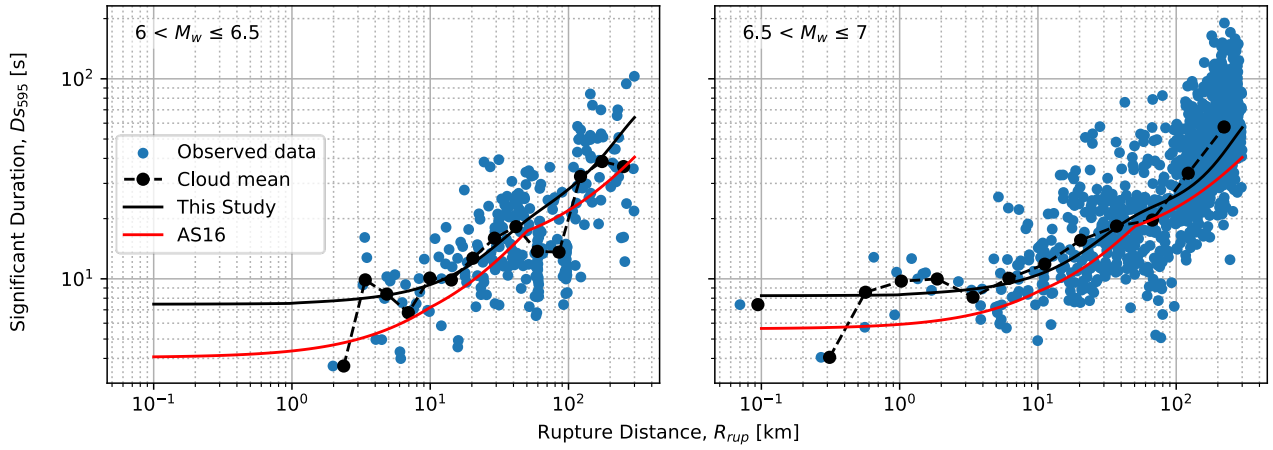


Figure 6. Distance attenuation plots of  $Ds_{595}$  for two different magnitude bins

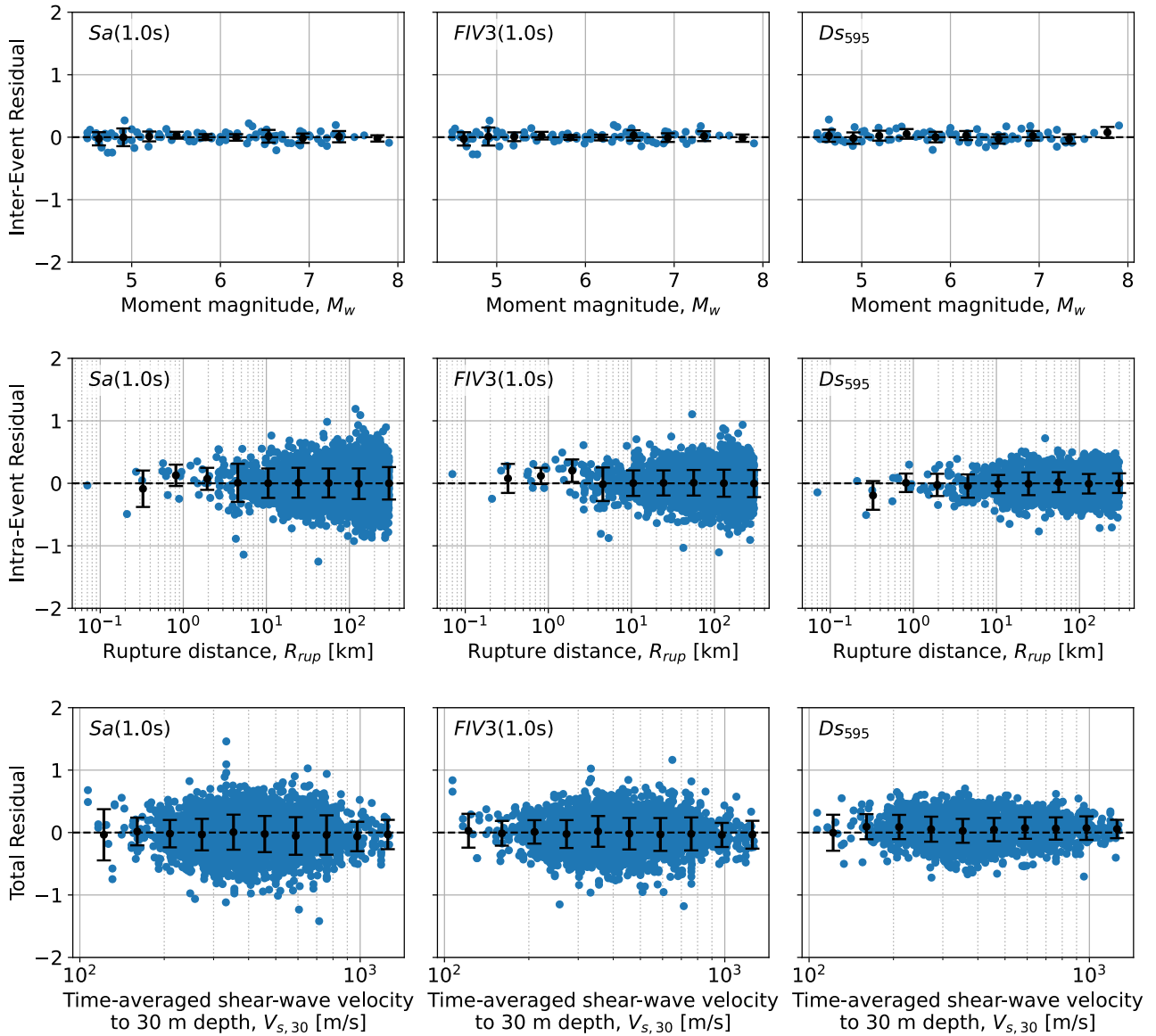


Figure 7. Inter-, intra-event, and total residuals versus  $M_w$ ,  $R_{rup}$ , and  $V_{s,30}$ , respectively, for three different IMs. Black dots and error bars represent the binned mean and  $\pm$  one standard deviation, respectively

To check for potential bias included in the model, the inter-, intra-event, and total residuals against three ground motion causal parameters (i.e.,  $M_w$ ,  $R_{rup}$  and  $V_{s,30}$ ) for three IMs (i.e.,  $Sa(1.0s)$ ,  $FIV3(1.0s)$ ,  $DS_{595}$ ) are plotted in Figure 7. It can be observed that there is no notable bias in the binned mean of residuals. Also, there is no significant change in standard deviations versus the GM causal parameters, which further corroborates the homoscedasticity assumption for the dispersion model of the proposed GGMM.

## 5.2. Standard deviations

The residuals between the values estimated from the model and those observed from the recorded ground motions were used to calculate the inter- and intra-event logarithmic standard deviations. The final standard deviations of all the IMs included in this study are presented in Figure 8, along with their counterparts given in GMMs from the literature for relative comparison. All standard deviations were transformed into natural logarithm (i.e.,  $\ln$ ) units to have an equal basis for comparison since the fitted GGMM was in terms of log base 10. It can be seen that the total standard deviation of the GGMM is the lowest for most IMs compared to that obtained from other GMMs available in the literature. This is the case especially for long period IMs. Additionally, the proposed model maintains a low inter-event standard deviation and about constant throughout all IMs. The high difference between intra- and inter-event standard deviations in this model, is because of the better characterisation of source effects, in comparison to path and site effects.

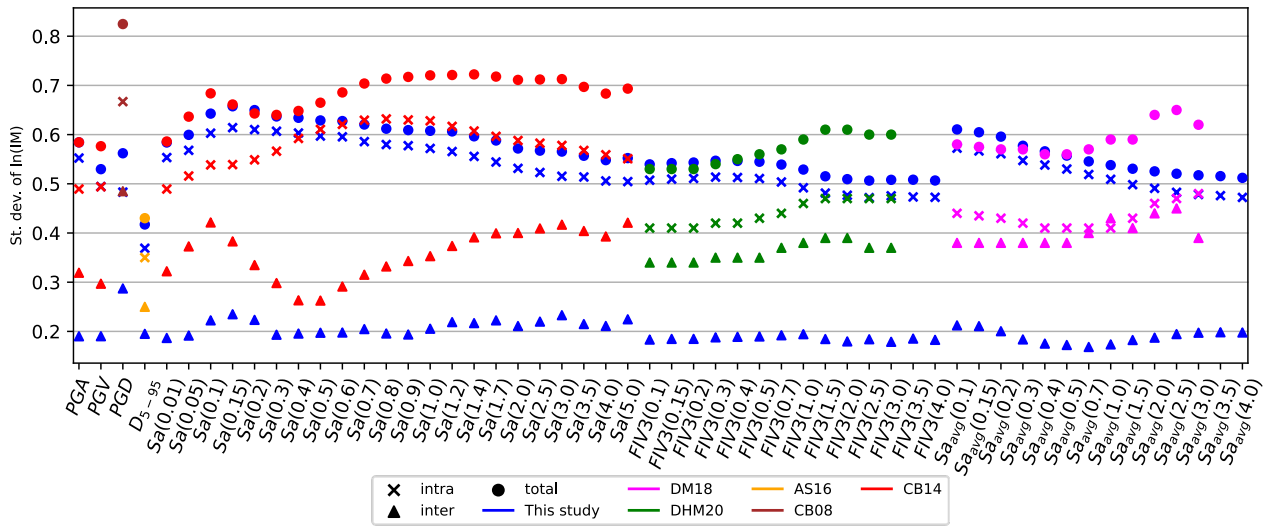


Figure 8. Inter-, intra-event and total standard deviations of the proposed model for all IMs, compared with models from the literature.

## 6. Correlation modelling

As previously stated, this GGMM which includes several IMs finds good utility in creating consistent (i.e., from same database and GMM) correlation models. Correlation models between  $Sa$ ,  $FIV3$  and  $DS_{595}$  were developed using this GGMM and presented in Aristeidou et al. (2024). As an example, the correlation coefficients between  $Sa$  and  $FIV3$  are illustrated in Figure 9.

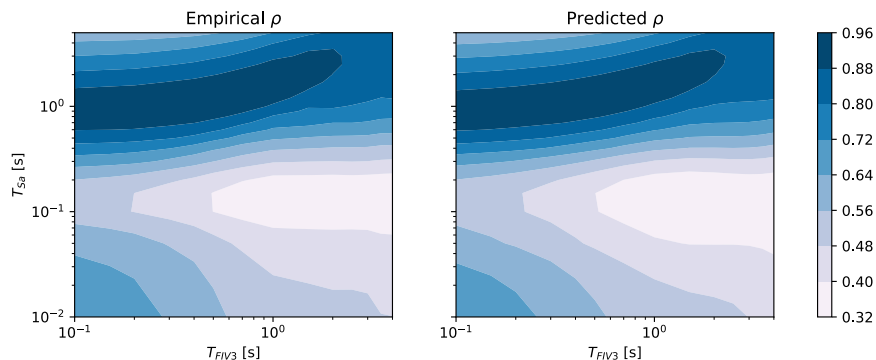


Figure 9. Empirical and corresponding predicted correlation coefficients between  $Sa$  and  $FIV3$

## 7. Summary and conclusions

This study proposed a non-parametric GGMM to estimate different types of amplitude and cumulative-based IMs (i.e.,  $PGA$ ,  $PGV$ ,  $PGD$ ,  $DS_{595}$ ,  $Sa(T)$ ,  $FIV3(T)$ ,  $Sa_{avg}(T)$ ). To demonstrate the potential of this ground motion modelling methodology, a total of 54 IMs were presented here. Nine ground motion causal parameters were used as an input to the non-parametric model, which describe the physics of source, path, and site characteristics sufficiently. This model was developed using artificial neural networks and was trained and tested using a stringently filtered subset of records from the NGA-West2 strong motion database. The variability was characterised by a homoscedastic standard deviation model. Thorough validation exercises and comparisons with other GMMs were carried out to demonstrate the suitability of the GGMM. It shows how this framework can effectively capture the complex relationships and interactions between different intensity measures and is one of the advantages of this GGMM, as it estimates various IMs in a single model. This helps develop more consistent correlation models between the estimated IMs since they come from the same database and GMM. Another advantage is that it minimises the dispersion of residuals (epistemic uncertainty), while keeping the two fitting performance metrics (i.e.,  $R^2$  and  $MSE$ ) at an optimal level. The logarithmic total standard deviations were low, especially in long-period IMs. All in all, the results and comparisons suggest that the model presents excellent performance in estimating a variety of IMs.

## 8. Acknowledgments

The work presented in this paper has been developed within the framework of the project “Dipartimenti di Eccellenza 2023-2027”, funded by the Italian Ministry of Education, University and Research at IUSS Pavia.

## 9. References

- Abadi M., Agarwal A., Barham P., Brevdo E., Chen Z., Citro C., Corrado G.S., Davis A., Dean J., Devin M., Ghemawat S., Goodfellow I., Harp A., Irving G., Isard M., Jia Y., Jozefowicz R., Kaiser L., Kudlur M., Levenberg J., Mané D., Monga R., Moore S., Murray D., Olah C., Schuster M., Shlens J., Steiner B., Sutskever I., Talwar K., Tucker P., Vanhoucke V., Vasudevan V., Viégas F., Vinyals O., Warden P., Wattenberg M., Wicke M., Yu Y., Zheng X. (2015). TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems,
- Abrahamson N.A., Youngs R.R. (1992). A stable algorithm for regression analyses using the random effects model, *Bulletin of the Seismological Society of America*, 82(1): 505–510.
- Afshari K., Stewart J.P. (2016). Physically parameterized prediction equations for significant duration in active crustal regions, *Earthquake Spectra*, 32(4): 2057–2081.
- Ancheta T., Darragh R., Stewart J., Seyhan E., Silva W., Chiou B., Wooddell K., Graves R., Kottke A., Boore D., Kishida T., Donahue J. (2013). PEER NGA-West2 Database, Technical Report PEER 2013/03,
- Aristeidou S., Shahnazaryan D., O'Reilly G.J. (2024). Correlation models for filtered incremental velocity with spectral acceleration and significant duration, *In: Proceedings of 18th world conference on earthquake engineering*. Milan, Italy.
- Atik L.A., Abrahamson N., Bommer J.J., Scherbaum F., Cotton F., Kuehn N. (2010). The Variability of Ground-Motion Prediction Models and Its Components, *Seismological Research Letters*, 81(5): 794–801.
- 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.
- Boore D.M., Stewart J.P., Seyhan E., 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.
- Bradley B.A. (2011). Empirical equations for the prediction of displacement spectrum intensity and its correlation with other intensity measures, *Soil Dynamics and Earthquake Engineering*, 31(8): 1182–1191.
- Campbell K.W., Bozorgnia Y. (2019). Ground Motion Models for the Horizontal Components of Arias Intensity (AI) and Cumulative Absolute Velocity (CAV) Using the NGA-West2 Database, *Earthquake Spectra*, 35(3): 1289–1310.
- Campbell K.W., Bozorgnia Y. (2014). NGA-West2 Ground Motion Model for the Average Horizontal Components of PGA, PGV, and 5% Damped Linear Acceleration Response Spectra, *Earthquake Spectra*, 30(3): 1087–1115.

- Campbell K.W., Bozorgnia Y. (2008). NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.01 to 10 s, *Earthquake Spectra*, 24(1): 139–171.
- Dávalos H., Heresi P., Miranda E. (2020). A ground motion prediction equation for filtered incremental velocity, FIV3, *Soil Dynamics and Earthquake Engineering*, 139: 106346.
- Dávalos H., Miranda E. (2018). A Ground Motion Prediction Model for Average Spectral Acceleration, *Journal of Earthquake Engineering*, 25(2): 319–342.
- Dávalos H., Miranda E. (2019). Filtered incremental velocity: A novel approach in intensity measures for seismic collapse estimation, *Earthquake Engineering and Structural Dynamics*, 48(12): 1384–1405.
- Derras B., Bard P.Y., Cotton F. (2014). Towards fully data driven ground-motion prediction models for Europe, *Bulletin of Earthquake Engineering*, 12(1): 495–516.
- Dhanya J., Raghukanth S.T.G. (2018). Ground Motion Prediction Model Using Artificial Neural Network, *Pure and Applied Geophysics*, 175(3): 1035–1064.
- Eads L., Miranda E. (2013). *Seismic collapse risk assessment of buildings: effects of intensity measure selection and computational approach*. Stanford, California: Stanford University.
- Eads L., Miranda E., Lignos D.G. (2015). Average spectral acceleration as an intensity measure for collapse risk assessment, *Earthquake Engineering & Structural Dynamics*, 44(12): 2057–2073.
- Fayaz J., Xiang Y., Zareian F. (2020). Generalized ground motion prediction model using hybrid recurrent neural network, *Earthquake Engineering and Structural Dynamics*, 50(6): 1539–1561.
- Haykin S.S. (2009). *Neural networks and learning machines*, New Jersey: Pearson.
- Kaklamanos J., Baise L.G., Boore D.M. (2011). Estimating unknown input parameters when implementing the NGA ground-motion prediction equations in engineering practice, *Earthquake Spectra*, 27(4): 1219–1235.
- Kiefer J., Wolfowitz J. (1952). Stochastic Estimation of the Maximum of a Regression Function, *The Annals of Mathematical Statistics*, 23(3): 462–466.
- Kingma D.P., Ba J. (2014). Adam: A Method for Stochastic Optimization *In: 3rd International Conference on Learning Representations, ICLR 2015* [Online]., 1–15. Available from: <http://arxiv.org/abs/1412.6980>.
- McCulloch W.S., Pitts W. (1943). A logical calculus of the ideas immanent in nervous activity, *The Bulletin of Mathematical Biophysics*, 5(4): 115–133.
- Močkus J. (1975). On bayesian methods for seeking the extremum *In:*, 400–404. Available from: [http://link.springer.com/10.1007/3-540-07165-2\\_55](http://link.springer.com/10.1007/3-540-07165-2_55).
- Picard R.R., Cook R.D. (1984). Cross-Validation of Regression Models, *Journal of the American Statistical Association*, 79(387): 575–583.
- Wooddell K.E., Abrahamson N.A. (2014). Classification of main shocks and aftershocks in the NGA-West2 database, *Earthquake Spectra*, 30(3): 1257–1267.