Peak Ground Acceleration Prediction using Artificial Neural Networks for Kachchh, Gujarat, India

Prantik Mandal (prantik@ngri.res.in)
NGRI https://orcid.org/0000-0002-5184-8584

Research Article

Keywords: Ground Motion Prediction, Artificial Neural Networks, Peak Ground Acceleration, Attenuation, Hypocentral Distances, Magnitudes

Posted Date: March 17th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1142608/v1

Abstract

The uncertainty in the empirical ground motion prediction models (GMMs) for any region depends on several parameters. In the present work, we apply an artificial neural network (ANN) to design a GMM of peak ground acceleration (PGA) for Kachchh, Gujarat, India, utilizing independent input parameters viz., moment magnitudes, hypocentral distances, focal depths and site proxy (in terms of average seismic shear-wave velocity from the surface to a depth of 30m (V_{s30})). The study has been performed using a PGA dataset consisting of eight engineering seismoscope (SRR) records of the 2001 M_w7.7 Bhuj earthquake and 237 strong-motion records of 32 significant Bhuj aftershocks of M_w3.3-5.6 (during 2002-2008) with epicentral distances ranging from 1.0 to 288 km.

We apply a feed-forward back propagation ANN method with 8 hidden nodes, which is found to be optimal for the selected PGA database and input-output mapping. The standard deviation of the error has been utilized to examine the performance of our model. We also test the ground motion predictability of our ANN model using real recordings of the 2001 Bhuj mainshock, two M_w5.6 Kachchh aftershocks and the 1999 M_w6.4 Chamoli mainshock. The standard deviation of PGA prediction error estimates in log10 units is found to be ±0.2554. Also, the model predictability of our ANN model suggests a good prediction of the PGA for earthquakes of M_w5.6-7.7, which are occurring in Kachchh, Gujarat, India.

Introduction

In general, ground motion prediction equations (GMPEs) or attenuation relationships form a key input for the earthquake hazard assessment of any region (Joyner and Boore, 1993; Ambraseys et al., 1996; Campbell, 1997). In general, the linear regression analysis is applied on the observed and/or synthetic peak ground acceleration data to determine GMPEs (Joyner and Boore, 1993; Ambraseys et al., 1996; Campbell, 1997; Boore and Atkinson, 2008; Nath et al., 2009; Anbazhagan et al., 2013). However, this method introduces errors in ground motion predictability due to the assumption of prior functional form. Further, the availability of less numbers of independent input parameters in India in comparison to other developed countries like USA and Europe has resulted in most of the available GPMEs in India as functions of magnitude and hypocentral distances (Nath et al., 2009; Anbazhagan et al., 2013). It is a well-established fact that inclusion of other independent input parameters (like focal depths, V_{s30}, soil conditions, focal mechanisms etc.) can lead to relatively robust prediction in ground motions (Abrahamson and Silva, 2008; Campbell and Bozorgnia, 2008; Castellaro et al., 2008; Kokusho and Sato, 2008). The CSIR-National Geophysical Research Institute (NGRI), Hyderabad, India, had monitored the earthquake activity in Kachchh during 2002 – 2016, with a close seismic network of 8-12 three-component broadband seismographs and 10-20 three-component strong-motion accelerographs (Fig. 1), which provided an excellent dataset for the robust estimation of moment magnitude, focal depths, hypocentral distances and focal mechanisms. Further toward hazard assessment, several micro-zonation studies have been carried in Kachchh, which provided V_{s30} estimates for the different regions of Kachchch that enabled us to construct maps of Vs at shallow depths below the Kachchh basin (Mandal

and Asano, 2019). Therefore, the availability of PGA data and input parameters (like Mw, focal depths, Hypocentral distances, and V_{s30}) for the Kachchh region has motivated us to apply an artificial neural network method to construct a ground motion model (GMM) for the PGA prediction.

Due to the availability of large data set, now-a-days the artificial neural network (ANN) has been used commonly to determine the GMPEs (Rumelhart et al., 2008; Derras et al., 2012 and 2016; Irshad et al., 2008). By applying an ANN methodology, a simple but powerful relationship is derived with Mw, Focal Depths, Hypocentral distances and shear velocity at 30 m depth (V_{s30}), for predicting PGA for Kachchh, Gujarat, India. For data fitting, pattern recognition and time series prediction, ANN has been demonstrated to be a powerful tool in a very complex situation. Massive parallel computations with a large number of variables can also be easily performed by an ANN, which can also grasp the non-linear aspects of the earthquake dataset. Thus, a data driven ANN can provide robust prediction for ground motions.

In the present work, we apply a feed-forward back propagation ANN method with 8 hidden layers, which is found to be optimal for the selected PGA database and input-output mapping for the Kachchh region. Here, the ANN method also uses the Levenberg-Marquardt training algorithm for mapping input-output datasets, available within ANN toolbox of MATLAB 2018b. Finally the ground motion predictability of the derived GMM has been tested using recorded PGAs of different Indian earthquakes and predicted PGAs from available GMPEs for India and eastern Northern America.

Artificial Neural Network (Ann) Method

Here, we have used feed-forward back propagation method consisting of an input layer, a hidden layer, and an output layer (Fig. 2a). One or more neurons (nodes) characterize each layer. In this method, first a collection of all nodes, which are multiplied with weights, is performed at each node of hidden layer. Next, a bias is attached to this sum, which is transformed through a non-linearity function prior to being transferred to the next layer. Functions such as hyperbolic tangent, sigmoid and linear functions can be used as transfer function. Subsequently, the same procedure is followed in this layer to obtain the network output results. The error between the network output and actual observation is estimated at the output layer. Finally, the desired outputs are achieved by back propagating this error to the input layer, through hidden layer in the network. All the synaptic weights of the network remain fixed during the forward processing while they are adjusted according to an error correction rule during the backward processing (Haykin, 1999). The FFBP network structure used in the present study is shown in Fig. 2b. Here, the input layer is considered to consist of x_n neurons while h_m neurons characterize the hidden layer. And, y_l neurons characterize the single output layer. The weights (w) consider to act as the link between very nodes in one layer to other layers. Here, w_{ij} marks the input to hidden weights while w_{ik} represents the hidden to output weights. And, the output of the network (y_k) (Gunaydin and Gunaydin, 2008) can be written as,

$$y_{ik} = \dot{f}[\sum_{j=0}^{m} w_{jk}.\, f[\sum_{i=1}^{n}(w_{ij}.\, x_i + b_1) + b_2] \text{ ————(1)}$$

where *b*1 and *b*2 are the biases for the first and second layer, respectively. The activation functions between input and hidden layers, and hidden and output layers are *f*(.) and $\dot{f}$(.), respectively.

Here, functions like the tangent sigmoid (tan-sgmoid), logarithmic sigmoid (log-sigmoid) and linear activation functions have been used for both *f*(.) and $\dot{f}$(.) for obtaining the best prediction performance.

For learning model, the back-propagation network model (Rumelhart, 1986) is used here, which aims at minimizing iteratively the mean sum squared error (MSE), that defined by

$E= \frac{1}{2P} \sum_{p=1}^{P} \sum_{k=1}^{l} (T_{pk} - y_{pk})$, p=1, 2, 3, ……P ----------(2)

where T_{pk} and y_{pk} are observed target and predicted output at k^{th} output node of p^{th} pattern, respectively. The total number of training patterns considered here is P. At each iteration, the global error (E) is minimized by adjusting the weights in each layer of the network until a convergence is achieved. Otherwise, the same process is repeated for further iterations.

A Learning Epoch is defined as each step in the learning phase. Here, for learning phase Levenberg-Marquardt algorithm (Levenberg, 1944; Marquardt, 1963) has been used that minimizes E and is expressed as:

$W_{new} = W_{old} - [J^T J + \gamma I]^{-1} J^T E (W_{old})$ ------(3)

where J is the Jacobian of the error function E, I is the identity matrix, and γ marks the iteration step value. Here, an adaptive learning rate is used that changes dynamically during the training stage from 0 to 1. We increase the learning rate by the factor learning increment if performance decreases toward the goal for an epoch. Otherwise, we adjust the learning rate by the factor learning decrement when performance increases for an epoch. Here, we use a value of 0.0001 as the performance goal throughout all FFBP simulations. After completing the training phase of the network successfully, a testing dataset (30% of the total data points) is used to examine the performance of the trained model.

The normalization of training input and output datasets are done through the following equation:

$x_{ni} = a. \frac{x_i - x_{min}}{x_{max} - x_{min}} + b$ ----------(4)

where x_i is the observed data obtained from i^{th} record and x_{ni} is the normalized value of the i^{th} record. And, x_{max} and x_{min} are the maximum and minimum values, respectively. Different values can be assigned for scaling factors a and b. Here, we have used different values for a, and b to obtain the best prediction performance. Finally different values of a and b are considered for different input and output parameters as discussed in the Supplementary data. This procedure has resulted in normalization of both inputs and output within the range of [-1.5, +1.5]. Several trials are made to select the optimum node numbers for the

hidden layers. Finally we obtain the best performance when we used 8 nodes for hidden layers. We stopped the networks training after maximum 10000 epochs.

Strong Motion Acceleration Dataset

For developing a GMM for PGA prediction using an ANN method, strong motion data of 237 recordings of 32 Kachchh earthquakes of M_w3.3-5.6, during 2002-2008, are used. For our study, eight engineering seismoscope (SRR) records of the 2001 M_w7.7 Bhuj earthquake are also utilized. The robust estimates of hypocentral parameters and moment magnitudes have been obtained by using data from 12 broadband seismographs and 10-20 strong-motion accelerogarphs. For our ANN modelling, we used the moment magnitudes (M_w), focal depths (D, m) and hypocentral distances (R, m) as independent inputs while we used V_{s30} (m/s) as the site proxy. The hypocentral distances for the dataset vary from 335 to 28900 m while focal depths and M_w are varying from 3-32 km and 3.3-7.7, respectively (Figs. 3a-d). The dataset is restricted to earthquakes in Kachchh, Gujarat for which the fault rupture lies mainly above a depth of 40 km. The values of V_{s30} have been obtained from several microzonation studies in Kachchh, Gujarat, which were used to construct three-dimensional shallow (0-0.9 km) shear velocity structure below the Kachchh basin (Mandal and Asano, 2019). For peak values we use the larger of the two horizontal components in the directions as originally recorded. Others (e.g., Campbell, 1981) have used the mean of the two components. The open circles on Figs. 3a-d show the distribution of the peak acceleration data with moment magnitudes, hypocentral distances, focal depths and V_{s30} values. For the 2001 M_w7.7 Bhuj earthquake the sources of strong-motion data have been given in previous publications (Cramer and Kumar, 2003; Iyengar and Raghukanth, 2004). The dataset consists of a total of 245 data points. For our ANN modelling, we used 70% (i.e. 171) of dataset for training while 30% of dataset has been used for testing and validation of the model.

Proposed Ann Model

Here, an artificial neural network (ANN) has been constructed to predict PGA wherein the site proxy $\log_{10}(V_{s30})$ has been used as an input. Other basic inputs are moment magnitude (M_w), $\log_{10}$(focal depth, D in m), and $\log_{10}$(hypocentral distance, R in m) while the output is considered as $\log_{10}$(y, in m/s*s). These inputs and output are normalized using equation 4 prior to applying the feed-forward ANN method. A feedforward supervised learning ANN architecture consisting of input, output, and a single hidden layer is chosen in this study as shown in Figs. 2a-b. One hidden layer is generally acknowledged for describing a nonlinear physical process. The input model with four input variables [Moment Magnitude, $\log_{10}$ (D), $\log_{10}(R)$, $\log_{10}(V_{s30})$] and eight number of nodes in the hidden layer which reported the least σ during three phases, is considered as the final selected model. The activation functions selected between the input-hidden layer and the hidden-output layer are Tanhsigmoid and linear, respectively. The ANN architecture of the derived model is shown in Fig. 2b.

In the analysis, the original data are rescaled to dimensionless values before being input to the networks (see the normalization of data section in the Appendix). We examined several values of number of nodes for hidden layers and finally selected 8 nodes that showed smallest validation loss among the tested models. The logistic sigmoid and rectified linear functions are used as activation functions in the last hidden layer and the other hidden layers, respectively. The linear (identity) function is used in the output layer.

Results And Discussions

Figure 4a shows the performance of our modelling suggesting that our model was trained for 11 times or epochs. However, by 5 iterations the model was trained and training was stopped because the model was overfitting after 5 iterations (Fig. 4a). Here we used the Levenberg-Marquardt training algorithm to train our dataset. The best validation performance is 0.08825 at epoch 5. The validation dataset of our model aligns with the best fit between 5 and 11 epochs while our model for training and test datasets suggest lower values than the best fit values between 2 and 11 epochs. Fig. 4b shows the error histogram for our NN modelling showing most of the larger positive and negative values are near to the zero error line (shown by yellow thick line), which suggests the error distribution of our NN modelling is good. Thus, our model is well trained.

Figures 5a-d show the regression analysis of the neural network modelling for our training, validation, test and complete datasets. Our training dataset consists of 171 data points (i.e. 70% of total data points) while validation and test dataset is having 37 data points (i.e. 15%) each. And, the complete dataset consists of 1 total of 245 data points. Our observed and predicted PGA values (through FFBP modelling) show high correlation coefficients exceeding 0.90381, except for the validation dataset (correlation coefficient ~ 0.88198). The all, training, and test datasets show a high correlation of 0.90381, 0.92177 and 0.90389, respectively (Figs. 5a, c, and d). Thus, apparently our NN modelling has trained and predicted the data quite well. From Figures 5a-d, we get four linear fit lines (equations 5-8) in the form $y = a\,x + b$, where x and y are observed and predicted (from the FFBP) PGA, respectively, which are mentioned below:

y = 0.83. x + 0.12 ——————————(5)

y = 0.76. x + 0.16 ——————————(6)

y = 0.84. x + 0.12 ——————————(7)

y = 0.82. x + 0.13 ——————————(8)

Figures 6a-b show good agreement between predicted and observed PGA values for training and test datasets while Figs. 6c-d show errors in predicted PGA for training and test datasets. A standard deviation of ±0.2554 is obtained for the errors in PGA prediction for training dataset while a standard

deviation of ±0.2720 is estimated for the errors in PGA prediction for the test data. Therefore, Figs. 6a-d suggest that the PGA prediction from our NN modelling is good.

Prediction of PGA and correlation with the observed PGA from moderate to large Indian Earthquakes

The predicted PGA estimates from our trained ANN model suggest a good agreement with the available observed PGA data for the 2001 Bhuj mainshock (derived from engineering seismographs records) and two 2006 Bhuj aftershocks of M_w5.6 (Figs. 7a-c). Since we do not have observed PGA data from Kachchh events of M6.0-6.5 events, thus, to validate the PGA predictions from our ANN model we have used available observed PGA data for the 1999 Chamoli earthquake of M_w6.4 (Fig. 7b). The predicted PGA values from our ANN model (thick red line in Fig. 7a) show an excellent agreement with the observed PGA data of the 7 March M_w5.6 2006 and 6 April M_w5.6 2006 Bhuj events. We also plotted the attenuation curve (thick grey line in Fig. 7a) predicated from the ground motion relation of Mandal et al. (2009) that does not show good correlation with the Observed PGA data of two M_w5.6 Bhuj events. However, the observed PGA data of the 1999 M_w6.4 Chamoli earthquake fall between the predicted PGA curves for M_w6.0 and M_w6.5 from our ANN model (Fig. 7b). This difference in PGA values could be attributed to the different source processes and tectonics for intraplate Kachchh and interplate Chamoli seismic zones. We also plotted the attenuation curve (thick grey line in Fig. 7b) predicated from the ground motion relation of Mandal et al. (2009) that does not show good correlation with the Observed PGA data of the 1999 Chamoli mainshock of M_w6.4. Since there is no observed PGA values are available for the 2001 Bhuj mainshock except one at 238 km away from the mainshock source zone (i.e. at Ahmedabad). Thus, we have used approximate corrected PGAs for the 2001 M_w7.7 Bhuj mainshock, which were derived from the Engineering Seismocope (SRR) records (Cramer and Kumar, 2003). Because of this fact, our ANN predicted PGA values for an M_w7.7 event show a fair but not very good agreement with the approximate PGA values derived from the SRR data of the 2001 mainshock (Fig. 7c). We also plotted attenuation curves for an M_w7.7 event predicted by available ground-motion relations for the eastern North America (Frankel et al., 1996; Toro et al., 1997; Somerville et al., 2001) and Kachchh, Gujarat, India (Mandal et al., 2009). From Fig. 7c, it is quite apparent that the predicted ground motion curve for an M_w7.7 event from our ANN model (shown by thick red line) shows a relatively better agreement with the observed data of the 2001 Bhuj earthquake as compared to other predicted curves from other attenuation relations. Thus, we can infer that the predictive PGA estimates from our ANN model suggest a fair to excellent correlation with the observed PGA values of earthquakes of M_w ranging from 5.6 to 7.7, which are occurring in Kachchh, Gujarat, India.

Conclusions

We derive a data driven feed-forward back propagation ANN model (with 8 hidden nodes) for predicting peak ground motions using observed PGA dataset of earthquakes occurring in Kachchh, Gujarat. The

dataset consists of thirteen engineering seismoscope (SRR) records of the 2001 M_w7.7 Bhuj earthquake and 232 strong-motion records of 32 significant Bhuj aftershocks of M_w3.3-5.6 (during 2002-2008) with epicentral distances ranging from 3.3 to 289 km. Our ANN predictive PGA values are validated by correlating with the observed PGA values of Kachchh events of M_w5.6 and 7.7, and 1999 Chamoli earthquake of M_w6.4. The standard deviation of PGA prediction error estimates in log10 units is found to be ±0.2554. Our derived ANN model for the Kachchh, Gujarat, India, is applicable within the range of Mw [3.3-7.7], focal depths [3-32 km], epicentral distances [1-288 km] and Vs30 [328-760 m/s].

Declarations

DATA AND RESOURCES

Most of the data that support the findings of this study are available in the supplementary material. Figures were plotted using the Generic Mapping Tool (GMT) software (Wessel et al., 2019; https://doi.org/10.1029/2019GC008515). All software and support data related to GMT software are freely accessible and available from this site (www.generic-mapping-tools.org). The elevation data used in generating GMT plots are obtained from the open source Digital Elevation Model (DEM) (https://asterweb.jpl.nasa.gov/gdem.asp). The artificial neural network (ANN) modelling code (NNSTART) in-built in Matlab software developed by MathWorks Inc. is used here.

DECLARATION OF COMPETING INTERESTS

The authors acknowledge there are no conflicts of interest recorded.

Acknowledgements

Authors are grateful to the Director, Council of Scientific and Industrial Research - National Geophysical Research Institute (CSIR-NGRI), Hyderabad, India, for his support and permission to publish this work. Authors are grateful to Matlab software of MathWorks Inc. for providing the in-built NNSTART script for the artificial neural network (ANN) modelling. The matlab script used for our modelling has been provided in the Supplementary data.

References

1. Abrahamson N, Silva W (2008) Summary of the Abrahamson & Silva NGA ground-motion relations. Earthq Spectra 24(1):67–97
2. Ambraseys NN, Simpson KA, Bommer JJ (1996) Prediction of horizontal response spectra in Europe. Earthquake Eng Struct Dynam 25(4):371–400
3. Anbazhagan P, Kumar A, Sitharam TG (2013) Soil Dyn Earthq Eng 53:92–108. https://doi.org/10.1016/j.soildyn.2013.06.003. Ground motion prediction equation considering

combined dataset of recorded and simulated ground motions

4. Boore DM, Atkinson GM (2008) Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s. Earthq Spectra 24(1):99–138. https://doi.org/10. 1193/1.2830434
5. Campbell KW (1997) Empirical near-source attenuation relationships for horizontal and vertical components of peak ground acceleration, peak ground velocity, and pseudo-absolute acceleration response spectra. Seismol Res Lett 68(1):154–179
6. Campbell KW, 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. Earthq Spectra 24(1):39–171
7. Castellaro S, Mulargia F, Rossi PL (2008) VS30: Proxy for seismic amplification? Seismol Res Lett 79(4):540–543
8. Derras B, Bard PY, Cotton F (2016) Site-conditions proxies, ground-motion variability and data-driven GMPEs: Insights from NGA-West 2 and RESORCE data sets. Earthq Spectra 32(4):2027–2056
9. Derras B, Bard PY, Cotton F, Bekkouche A (2012) Adapting the neural network approach to PGA prediction: An example based on the KiK-net data. Bull Seismol Soc Am 102(4):1446–1461
10. Frankel A, Mueller C, Barnhard T, Perkins D, Leyendecker EV, Dickman N, Hanson S, Hopper M (1996) National seismic hazard maps: Documentation June 1996, Open-File Report (United States Geological Survey) 96–532, 41 pp
11. Gunaydın K, Gunaydın A (2008) Peak Ground Acceleration Prediction by Artificial Neural Networks for Northwestern Turkey, Hindawi Publishing Corporation Mathematical Problems in Engineering Volume 2008. Article ID 919420:20. doi:10.1155/2008/919420
12. Haykin S (1999) Neural Networks: A Comprehensive Foundation, Prentice-Hall, Englewood Cliffs, NJ, USA, 2nd edition
13. Irshad A, Naggar MHEI, Khan AN (2008) Neural network based attenuation of strong motion peaks in Europe. J Earthq Eng 12:663–680
14. Joyner WB, Boore DM (1993) Bull Seismol Soc Am 83(2):469–487Methods for regression analysis of strong-motion data
15. Kokusho T, Sato K (2008) Surface-to-Base amplification evaluated from KiK-net vertical array strong motion records. Soil Dynam Earthq Eng 28(9):707–716
16. Levenberg K (1944) A method for the solution of certain non-linear problems in least squares. Q Appl Math 2:164–168
17. Mandal P, Kumar N, Satyamurthy C, Raju IP (2009) Ground-motion Attenuation Relation from Strong-motion Records of the 2001 Mw 7.7 Bhuj Earthquake Sequence (2001–2006), Gujarat, India, Pure. Appl Geophys 166:1–19
18. Marquardt D (1963) An algorithm for least-squares estimation of nonlinear parameters. Journal of the Society for Industrial & Applied Mathematics 11:431–441

19. Nath SK, Raj A, Thingbaijam KK, Kumar A (2009) Ground motion synthesis and seismic scenario in Guwahati city—a stochastic approach. Seismol Res Lett 80(2):233–242. https://doi.org/10.1785/gssrl.80.2.233
20. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. Nature 323:533–536
21. Somerville P, Collins N, Abrahamson N, Graves R, Saikia C (2001) Ground motion attenuation relations for the central and eastern United States, Final report to the U.S. Geological Survey, 30 June 2001. URS Group, Inc., Pasadena, California, p 36
22. Toro G, Abrahamson N, Schneider J J (1997) Model of strong ground motions from earthquakes in central and eastern North America: Best estimates and uncertainties. Seismol Res Lett 68:41–67

Figures

Figure 1

Elevation (in m) map of the Kachchh, Gujarat, India, showing three-component strong-motion stations (marked by solid filled blue triangles). Black dotted lines mark different tectonic faults in the region viz., ABF: Allah Bund fault; IBF: Island Belt fault; GF: Gedi fault; NWF: North Wagad fault; KMF: Kachchh mainland fault; KHF: Katrol Hill fault. The inset shows the elevation map of India showing locations of the 1967 M_w6.3 Koyna, 1999 M_w6.4 Chamoli, and 2001 M_w7.7 Bhuj earthquakes. The study area is marked by a yellow square. AS marks the Arabian sea and BOB represents the Bay of Bengal.

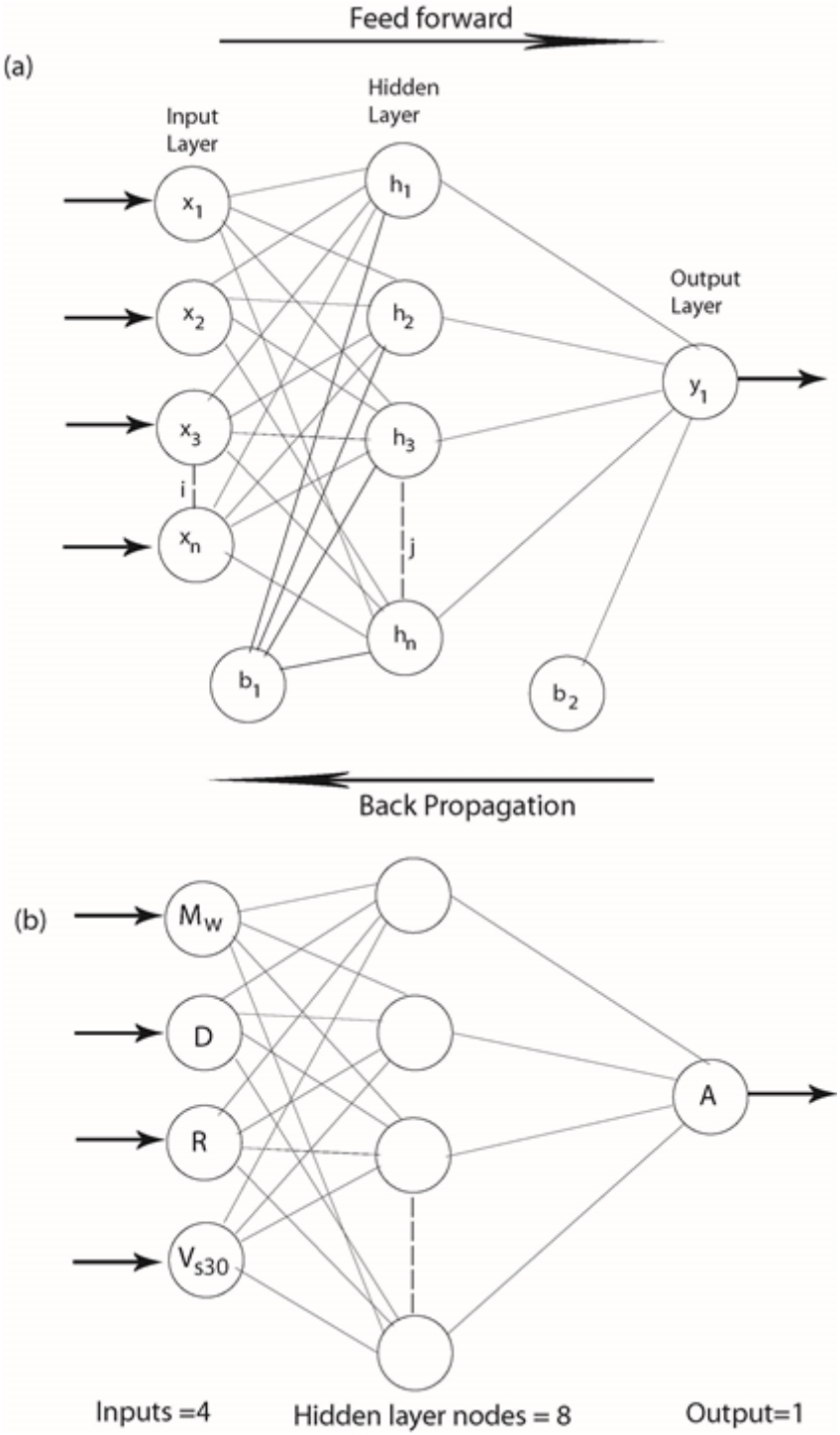

Figure 2

(a) The schematic diagram of a feed-forward back propagation artificial neural network (ANN) showing weights, biases, input layers, hidden and output layers. (b) Structure of the NN used in this study which uses V_{s30} site proxies as input variables. The NN network consists of 4 inputs, 8 hidden layer nodes and one output layer.

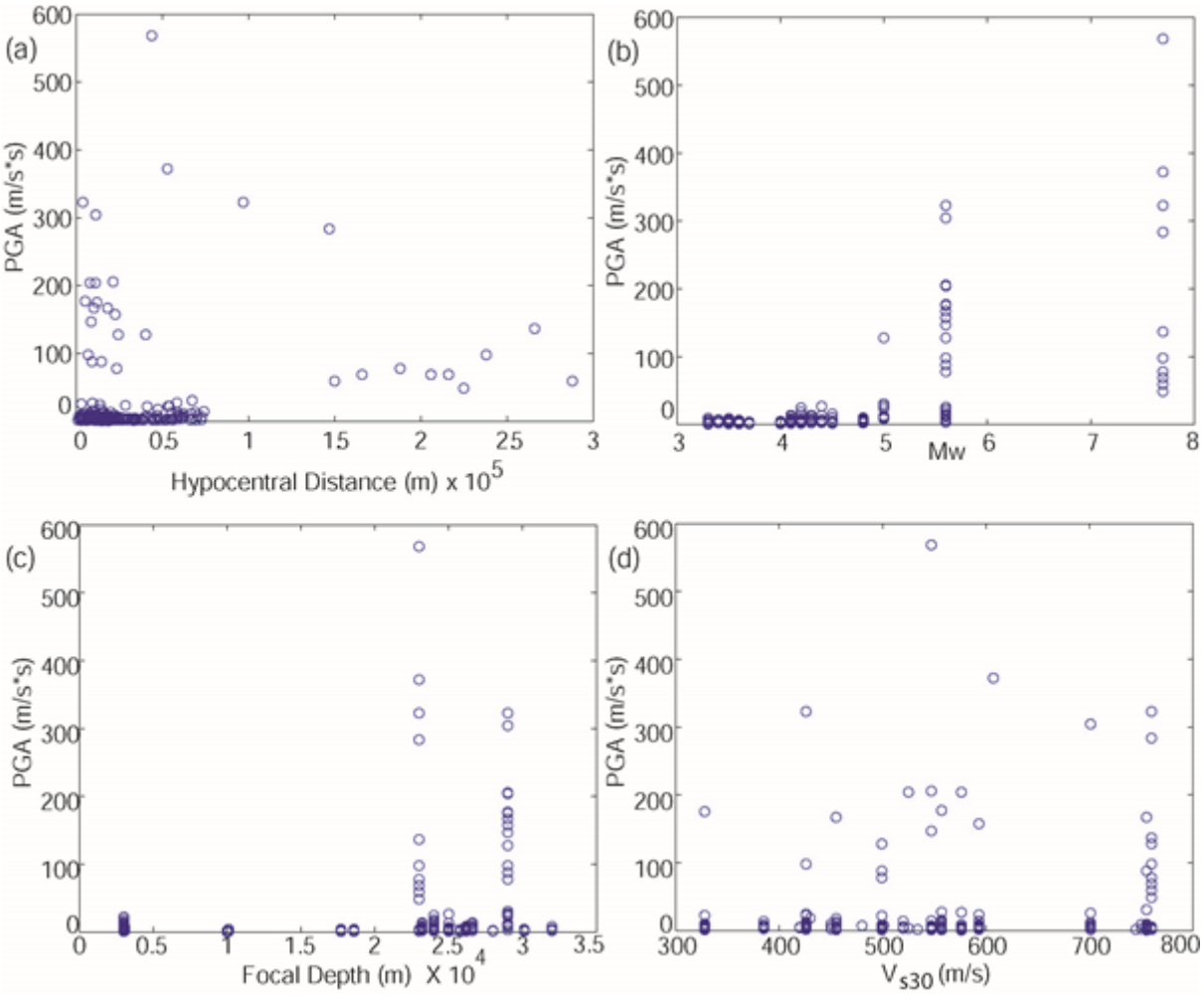

Figure 3

Observed data used in the ANN modelling. Distribution of records with respect to (a) Peak ground acceleration (PGA, in m/s*s) and hypocentral distances (in m), (b) PGA and moment magnitudes (M_w), (c) PGA and Focal depths (in m), and (d) PGA and V_{s30} (in m/s).

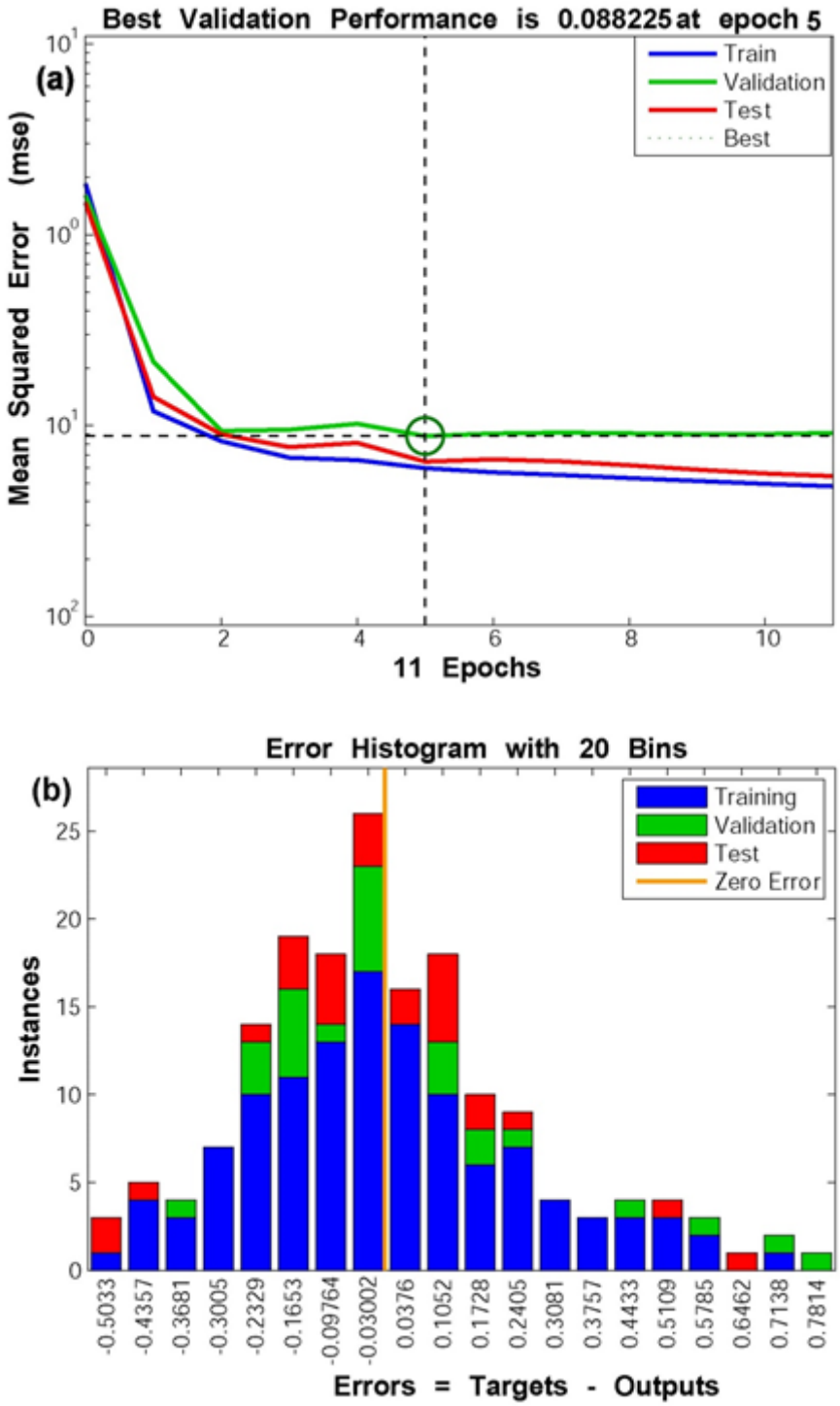

Figure 4

(a) A plot between Mean squared error (MSE) and numbers of Epochs showing the validation of the best performance of the FFBP ANN used in the study is 0.08825 at epoch 5. The performance of ANN on the training, validation and test data is shown by blue, green and red lines, respectively, while the best performance is marked by grey dotted line, (b) A plot showing error histogram for the training, validation and test data.

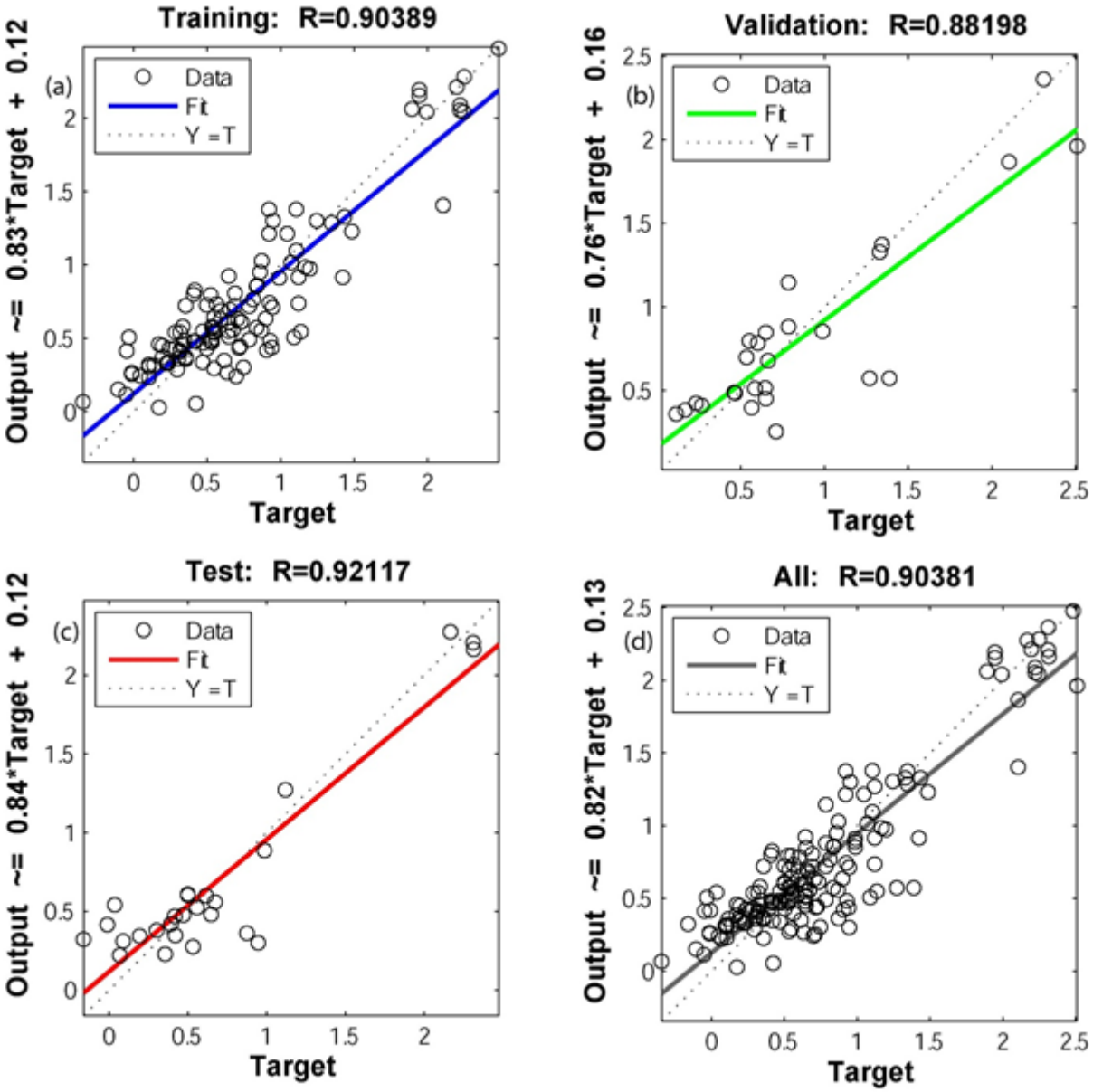

Figure 5

Regression analysis between the output and target of the NN modelling (a) Training (70% of all the dataset points) showing correlation R=0.90389, (b) Validation (15% of all the data points) showing correlation R=0.88198, (c) Testing (15% of all the data points) showing correlation R=0.92117, and (d) all Data points showing correlation R=0.90381.

Figure 6

(a) Predicted $\log_{10}$(PGA) from the NN modelling for training dataset, (b) Predicted $\log_{10}$(PGA) from the NN modelling for testing dataset, (c) Error in Predicted $\log_{10}$(PGA) from the NN modelling for training

dataset showing a standard deviation of ±0.2554, and (d) Error in Predicted $\log_{10}$(PGA) from the NN modelling for testing dataset showing a standard deviation of ±0.2720.

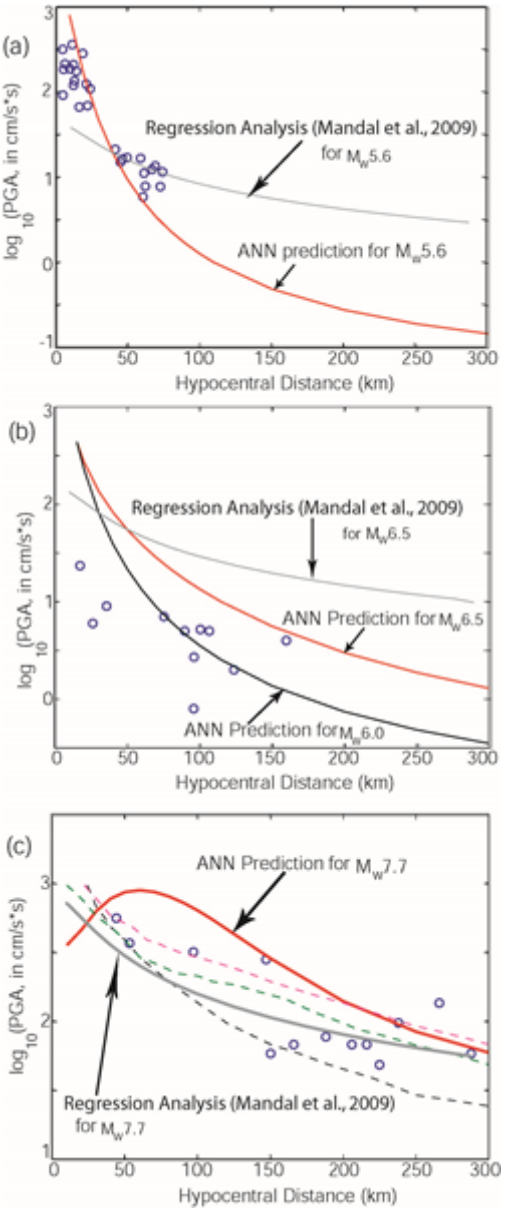

Figure 7

The attenuation curves (solid thick red line) for the predicted PGA using the feed-forward data-driven ANN trained network for (a) M_w5.6, (b) M_w6.5 and (c) Mw7.7 events. Observed PGA data (blue open circles) for the 2001 Bhuj mainshock and large Bhuj aftershocks of M_w5.6 and the 1999 Chamoli earthquake have also been plotted to demonstrate the robustness of the prediction. In Fig. 7b, black solid line marks the attenuation curve for the predicted PGA (using ANN model) for an M_w6.0 earthquake. In Fig 7c, predicted curves for the ground motion attenuation models of Frankel et al. (1996) (dotted green line), Somerville et

al. (2001) (dotted magneta line), Toro et al. (1997) (dotted brown line), and Mandal et al. (2009) (dotted brown line).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- nhazSupplementaryInformation.docx