Computers and Structures 89 (2011) 2176–2194

Contents lists available at SciVerse ScienceDirect

Computers and Structures

journal homepage: www.elsevier.com/locate/compstruc

Prediction of principal ground-motion parameters using a hybrid method coupling artificial neural networks and simulated annealing

Amir Hossein Alavi [a], Amir Hossein Gandomi [b,*]

[a] School of Civil Engineering, Iran University of Science and Technology, Tehran, Iran
[b] Department of Civil Engineering, The University of Akron, Akron, OH 44325, USA

ARTICLE INFO

Article history:
Received 10 November 2010
Accepted 26 August 2011
Available online 6 October 2011

Keywords:
Time-domain ground-motion parameters
Artificial neural networks
Simulated annealing
Attenuation relationship
Nonlinear modeling

ABSTRACT

In this study, new models are derived to predict the peak time-domain characteristics of strong ground-motions utilizing a novel hybrid method coupling artificial neural network (ANN) and simulated annealing (SA), called ANN/SA. The principal ground-motion parameters formulated are peak ground acceleration (PGA), peak ground velocity (PGV) and peak ground displacement (PGD). The proposed models relate PGA, PGV and PGD to earthquake magnitude, earthquake source to site distance, average shear-wave velocity, and faulting mechanisms. A database of strong ground-motion recordings released by Pacific Earthquake Engineering Research Center (PEER) is used to establish the models. For more validity verification, the ANN/SA models are employed to predict the ground-motion parameters of a part of the database beyond the training data domain. ANN and multiple linear regression analyses are performed to benchmark the proposed models. Contributions of the input parameters to the prediction of PGA, PGV and PGD are evaluated through a sensitivity analysis. The ANN/SA attenuation models give precise estimations of the site ground-motion parameters. The proposed models perform superior than the single ANN, regression and existing attenuation models. The optimal ANN/SA models are subsequently converted into tractable design equations. The derived equations can readily be used by designers as quick checks on solutions developed via more in-depth deterministic analyses.

1. Introduction and background

Seismic hazard analysis is an essential step in engineering phase. The seismological characteristics of earthquakes usually include magnitude, distance, faulting type, and soil effects. Time-domain and response-domain parameters are well-known engineering parameters of an earthquake. Three major parameters of the time-domain class are: (1) peak ground acceleration (PGA), (2) peak ground velocity (PGV), and (3) peak ground displacement (PGD) [1]. The time and response-domain parameters can both be applied to structural risk assessment. The spectral parameters are shown to be more efficient than the time-domain parameters [2]. On the other hand, the time-domain parameters are more practical due to their independency from structures. Thus, PGA, PGV and PGD are frequently used in seismic hazard studies. These key elements can be predicted using different methods such as on-site investigation and physical modeling. In most cases, implementing these methods is extensive, cumbersome and costly [1,3]. Much effort should be made to describe limited observations through the physical modeling of an earthquake. The physical models are usually developed in the context of stochastic modeling approach and random vibration theory [4]. More advanced physical modeling methods simulate the realistic process of faulting through the numerical analysis of crack and wave propagation [5].

An empirical approach to assess the ground-motion engineering parameters is to use attenuation relationships. The attenuation relationships play a key role in seismic hazard analysis. They often correlate the ground-motion parameters with various independent variables (e.g., earthquake magnitude, distance from source to site, local site conditions, and earthquake source characteristics) [1,3,6]. It is not an easy task to develop a correlation between PGA, PGV and PGD and the predictor variables due to high nonlinearity in the relationships. Regression analysis is a conventional way to build the attenuation relationships from the recorded strong motion data [7–10]. In this context, Fajfar and Perus [11] proposed a non-parametric multidimensional regression method for the prediction of the seismic ground-motion parameters. Perus and Fajfar [12] used a non-parametric approach, called conditional average estimator (CAE) method, for the ground-motion prediction. In addition to physical aspects [5,13], the commonly used regression analysis has major drawbacks related to the idealization of complex processes, approximation, and averaging widely varying prototype conditions. Also, the nature of the corresponding problem should be pre-defined by a linear or nonlinear equation to perform the regression analysis [1]. Thus, the derived attenuation models

* Corresponding author. Tel.: +1 234 788 0619.
E-mail addresses: ah_alavi@hotmail.com (A.H. Alavi), a.h.gandomi@gmail.com, ag72@zips.uakron.edu (A.H. Gandomi).

0045-7949/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compstruc.2011.08.019

are often limited in their ability to efficiently simulate the complex behavior of the ground-motion parameters [2]. The issues raised above suggest the necessity of utilizing more robust methods to predict the ground-motion parameters.

Empirical modeling by artificial intelligence techniques, such as genetic programming (GP) and artificial neural networks (ANNs), is a different approach to estimate the ground-motion characteristics. Such methods have a great capability of adaptively learning from experience and extracting various discriminators. Recently, Cevik and Cabalar [14] utilized a branch of GP, namely gene expression programming (GEP) to derive a greatly simplified prediction equation for PGA upon a strong ground-motion data from Turkey. Gandomi et al. [1] presented a hybrid method coupling GP and orthogonal least squares, called GP/OLS, to derive new ground-motion prediction equations. Kerh and Chu [15] employed ANNs to predict PGA at two main line sections of Kaohsiung Mass Rapid Transit in Taiwan. Chu et al. [16] developed an ANN model to analyze the strong motion characteristics around the Kaohsiung area of Taiwan. Kerh and Ting [17] used ANNs to predict PGA along a high-speed rail system in Taiwan. Gullu and Ercelebi [3] and Gunaydin and Gunaydin [18] developed prediction models for PGA using ANNs upon a strong motion database from Turkey. Ahmad et al. [19] established ANN-based attenuation relationships for PGA, PGV and PGD using the European earthquake data. A major constraint in application of ANN is the network's tendency to become trapped in local minima [20]. To cope with this problem and to obtain an optimal solution, a neural network may be trained using global search algorithms such as genetic algorithms [21,22], tabu search [23] and evolutionary strategies [24]. Simulated annealing (SA) has also been used by researchers [25–27] for training ANNs as an alternative to the more traditional local search algorithms (e.g., gradient search techniques). Recently, Ledesma et al. [28] combined ANNs and SA to make a hybrid algorithm with better efficiency. They proposed a novel cooling schedule based on temperature cycling for implementing SA to improve the ANN training. It was shown that the networks trained using temperature cycling outperformed those trained by the conventional exponential or linear cooling schedules [28]. Despite remarkable prediction capabilities of this hybrid ANN/SA method [28], there has not been yet any scientific efforts directed at applying it to civil engineering tasks.

In this study, the ANN/SA technique is utilized to derive new generalized attenuation relationships for PGA, PGV and PGD. The employed hybrid system uses the SA strategy to assign good starting values to the weights of the network before performing optimization. ANN/SA is useful in deriving prediction models for PGA, PGV and PGD by directly extracting the knowledge contained in the experimental data. ANN-based models are commonly considered as black-box systems as they are unable to explain the underlying principles of prediction. To overcome this limitation, the optimal ANN/SA models are converted into relatively simple design equations. A conventional calculation procedure is further proposed based on the fixed connection weights and bias factors of the best obtained structures. The predictions made by the developed models are further compared with those provided by the ANN, regression and empirical models [7,10,29]. The proposed models are developed based on a comprehensive database of strong ground-motions assembled by Pacific Earthquake Engineering Research Center (PEER) [30].

2. Methodology

2.1. Artificial neural network

ANNs have emerged as a result of simulation of biological nervous system. The ANN method was founded in the early 1940s by McCulloch and co-workers [31]. The first researches were focused on building simple neural networks to model simple logic functions. ANNs can be applied to a variety of problems without algorithmic solutions or problems with complex solutions. ANNs formulate a mathematical model for a system in which no clear relationship is available between inputs and outputs. Multilayer perceptron (MLP) network [32] is the most well-known class of ANNs. MLPs have feed-forward architectures. They are essentially capable of approximating any continuous function to an arbitrary degree of accuracy [32]. These networks are usually applied to perform supervised learning tasks, which involve iterative training methods to adjust the connection weights within the network [33]. They are usually trained with back-propagation algorithm. Fig. 1 shows a schematic representation of an MLP network. The MLP network consists of an input layer, at least one hidden layer of neurons and an output layer. Each of these layers has several processing units and each unit is fully interconnected with weighted connections to units in the subsequent layer. Each layer contains a number of nodes. Every input is multiplied by the interconnection weights of the nodes [33]. Finally, the output (h_j) is obtained by passing the sum of the product through an activation function as follows:

$$h_j = f\left(\sum_i x_i w_{ij} + b\right) \tag{1}$$

where $f()$ is activation function, x_i is the activation of ith hidden layer node, w_{ij} is the weight of the connection joining the jth neuron in a layer with the ith neuron in the previous layer, and b is the bias for the neuron. For nonlinear problems, the sigmoid functions (hyperbolic tangent sigmoid or log-sigmoid) are commonly adopted as the activation function. Adjusting the interconnections between layers will reduce the following error function [34,35]:

$$E = \frac{1}{2}\sum_n \sum_k \left(t_k^n h_k^n\right)^2 \tag{2}$$

where t_k^n and h_k^n are respectively the calculated output and the actual output value, n is the number of sample and k is the number of output nodes. Further details of MLPs can be found in [32].

2.2. Simulated annealing

SA is a general stochastic search algorithm used for solving optimization problems. This algorithm was first applied to optimization problems by Kirkpatrick et al. [36] and Cerny [37]. SA is very useful for solving several types of optimization problems with nonlinear functions and multiple local optima [38,39]. SA makes use of the Metropolis algorithm [38] for computer simulation of annealing. Annealing is a process in which a metal is heated to a high temperature and thereafter it is gradually cooled to relieve thermal stresses. During the cooling process, each atom takes a specific position in the metal crystalline structure [40]. By changing the

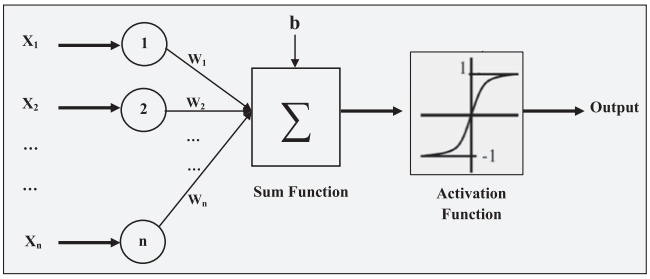

Fig. 1. A schematic representation of an MLP neural network.

temperature, the crystalline structure changes to a different configuration. An internal energy (E) can be measured and assigned to each state of the crystalline structure of the metal which is achieved during the annealing process. At each temperature within the annealing process, the atoms are allowed to adjust to a stable equilibrium state of least energy if the temperature does not decrease quickly [40]. It is obvious that changing the metal crystalline structure, through the annealing, is associated with a change in the internal energy as ΔE. However, as the metal temperature gradually drops down, the overall trend of the internal energy change follows a decreasing process, but sometimes the energy may increase by chance. The metal crystalline structure achieves near global minimum energy during the process of annealing. This process is simulated by SA to find the minimum of a function in a certain design space. The objective function corresponds to the energy state and moving to any new set of design variables corresponds to a change of the crystalline structural state [40]. The abilities and shortcomings of SA are well summarized by Ingber [41].

2.3. Hybrid artificial neural network-simulated annealing

ANNs can be used to learn from a set of data known as training set. During the training process, the network's weights are adjusted until the actual and predicted output values become as close as possible [42]. It is reasonable to re-design the network structure to achieve a specified performance. Once the network has been trained, it must be evaluated using another set, called testing set. The training procedure is typically divided into two main steps: (1) initialization and (2) optimization. The initialization process might involve assigning initial random values to the weights of the network [43]. This process might be more sophisticated by using powerful methods such as SA for assigning good starting values to the weights. The optimization process is often a gradient-based algorithm. This process requires a good starting point to be successful. Thus, a complete training process depends on both the initialization process and the optimization process [28].

In contrast with other optimization methods, SA is a no greedy optimization approach. Thus, it does not fall easily into local minima [28]. Using a high quality random generator is one of the most important practical considerations when implementing the SA algorithm [28,44]. For the ANN training, SA randomly perturbs the weights of the network following a cooling schedule. Once the weights are perturbed, the network performance is evaluated using an appropriate training set. The cooling schedule in SA is generally linear or exponential. In addition, the cooling schedule may iterate at each temperature or may increase the number of iterations at a specific temperature when improvement occurs [45]. Each time a new solution is perturbed and evaluated, the specialist decides whether the new solution is accepted or rejected using the Metropolis algorithm [28,44]. Some implementations of SA accept a new solution only if the new solution is better than the old one. That is, SA accepts the solution only when the error decreases [46]. It is often more efficient to accept the solution even when its corresponding error is less than that of the previous solution. The probability of accepting a new solution (P_a) was first incorporated into numerical calculations by Metropolis et al. [38] as given below:

$$P_a(\Delta E, y) = \begin{cases} e^{-\frac{k\Delta E}{y}} & \Delta E > 0 \\ 1 & \Delta E > 0 \end{cases} \tag{3}$$

where

ΔE = Error (new solution) − Error (current solution),

y = Current temperature,

and k is the acceptance constant depending on the range of the weights and inputs of the network. At high temperatures, the algorithm may frequently accept a solution even if it is not better than the previous one [47]. During this phase, the algorithm searches for an optimal solution in a very wide range with no concern about the quality of the solution. As the temperature decreases, the algorithm is more selective. In this case, it accepts a new solution only if its error is less than or very similar to the previous solution error [28].

An important factor when implementing SA for the ANN's initialization is to choose an appropriate acceptance constant, k. This constant affects the probability of acceptance for a given value of ΔE. The k value is dependent on the temperature range, training set and weight allowed values. An important issue during the preparation of the training set is to appropriately scale the data. Considering wide ranges for the inputs is not recommended for SA because a wider range means a wider range of weights. Therefore, more combinations are required for SA leading to increases in training time. Once the inputs and weights are limited, it is possible to choose an appropriate value for k to provide reasonable learning. A reasonable learning means that the error does not wander excessively from high and low values due to an excessive probability of acceptance (excessive heating) [28]. Typical cooling schedules start at high temperatures and gradually cool down using different functions until a specified final temperature is reached [48]. On the other hand, the hybrid ANN/SA method requires the temperature to increase and decrease periodically. Fig. 2 shows typical linear and temperature cycling cooling schedules.

The linear schedule is very popular. The temperature cycling schedule has been previously utilized by Mobius et al. [49] to solve optimization problems. To use temperature cycling for the ANN learning, the finite length series ($x[n]$) is defined as:

$$x[n+1] = \rho x[n], \quad n = 1, 2, 3, \ldots, N \tag{4}$$

where N is the number of temperatures, $x[1]$ is the initial temperature, $x[N]$ is the final temperature, and ρ is a cooling constant expressed as:

$$\rho = e^{log}\left(\frac{(N-1)_x[N]}{x[1]}\right) \tag{5}$$

The cooling schedule for temperature cycling is:

$$y[n] = \sum_{m=0}^{M-1} x[n - mN] \tag{6}$$

where $y[n]$ is the SA temperature, M is the number of cycles before starting the optimization process. Note that the number of iterations at each temperature for temperature cycling is recommended to be a relatively low value. If training of the network is performed using too many iterations (e.g., 100) per temperature, the benefit gained for using temperature cycling is lost. This is because the solution may fall too much and it will be difficult to escape from this minimum [28]. In addition, the network's weights must be updated using a recursive equation as given below:

$$w_{i,j}[n+1] = \gamma(1-\lambda)w_{i,j}[n] = \lambda u[n] - \frac{1}{2}\lambda \tag{7}$$

where $w_{i,j}$ is the network's weight connecting the j neuron with the i neuron in the next layer, λ is the perturbation ratio defined as:

$$\lambda = \frac{y[n]}{x[1]} \tag{8}$$

$u \in [0,1)$ is a uniformly distributed random variable, and γ depends on the weights range. For a current temperature $y[n]$ close to the initial temperature $x[1]$, λ takes values close to 1 and the network's weights are violently perturbed. As the temperature decreases, the

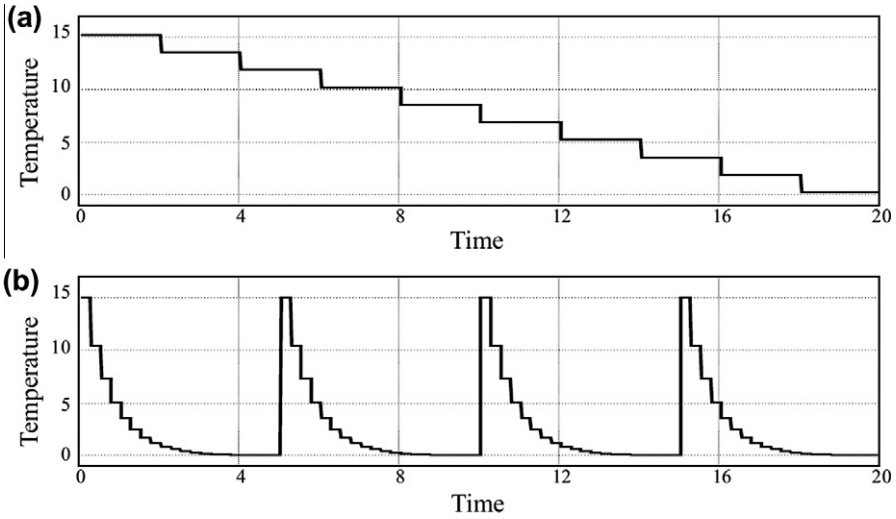

Fig. 2. Cooling schedule (a) linear cooling, (b) temperature cycling [28].

weights wandered around a fixed value at random. Also, Eq. (5) properly combines the perturbation and the current weight value. Other approaches for SA perturb the network's weights, and then clip the perturbed value to keep the weights within a specified range. This is not reasonable since some information is lost during the clipping process [28].

3. Modeling of time-domain strong ground-motion characteristics

The damage potential of an earthquake depends on the ground-motion characteristics and local site conditions. Amplitude, frequency content and duration of motion are the major characteristics of earthquake motion [1,6]. PGA, PGV and PGD have a key role in explaining the characteristics of strong ground-motions. These parameters are frequently presented as functions of different seismic variables. The most remarkable seismological aspects that influence the ground-motion parameters are the source effect, path effect and site effect [1]. The source effect is related to numerous parameters such as the level of stress drop in the earthquake event, mechanism of faulting, and direction of faulting. The path effect is related to the distance of the site from the fault. Different definitions are presented for the distance in the literature such as closest distance and Joyner–Boore distance [9]. The site effect is a significant element included in the attenuation relationships. Some of the models consider the site effect in a generic way (i.e. soil or rock) and the others uses the soil shear velocity as an indicator of the site effect [1]. However, modern attenuation relationships [8–10] are mainly in terms of the earthquake magnitude, source to site distance, geotechnical condition of site, and faulting mechanism. Other affecting physical parameters such as stress drop, rupture propagation, directivity, and nonlinear soil behavior notably reflect the uncertainties. Therefore, they are not usually used in the development of the predictive equations [1,14].

Advanced approaches have recently been employed to enhance the precision of the conventional regression-based analyses. The study uses the ANN/SA approach to derive alternative attenuation relationships for PGA, PGV and PGD. The most important factors representing the ground-motion parameters behavior are selected based on a trial study and after a literature review [1,3,5,7–10,14,50]. Consequently, the formulations of PGA (cm/s^2), PGV (cm/s) and PGD (cm) are considered to be as follows:

$$\begin{pmatrix} \ln(PGA) \\ \ln(PGV) \\ \ln(PGD) \end{pmatrix} = f(F, M_w, \ln(R_{ClstD}), V_{s30}) \tag{9}$$

where,

F: Indicator variable representing different fault types

1: Reverse (dip slip with hanging-wall side up)
2: Normal (dip slip with hanging-wall side down)
3: Strike-slip (horizontal slip)

M_w: Earthquake magnitude (moment magnitude)
R_{ClstD} (km): Closest distance to the rupture surface
V_{s30} (m/s): Average shear-wave velocity over the top 30 m of site

Indicator variables representing the style of faulting are defined in terms of rake angle. The rake angle is described as the average angle of slip in degrees measured in the plane of rupture between the strike direction and the slip vector [10]. Campbell and Bozorgnia [10] stated that directly using the indicator variables for the style of faulting rather than the rake angle may lead to better results.

3.1. Strong-motion database and data preprocessing

A database compiled by Power et al. [30] (NGA Flatfile V 7.3) in the PEER-NGA project is employed for the model development. The database is comprised of shallow crustal earthquakes recorded data at active tectonic regions of the world. The database covers a broad range of magnitude and distance. In this study, a part of the database is excluded from the analysis considering some of the filtering strategies proposed by Boore and Atkinson [10]. Data sets missing the required information and also the duplicate records are omitted. Finally, out of the total of 3551 records, 2815 records for three different fault types are employed for the model development. The predictor variables included in the analysis are F, M_w, R_{ClstD} (km), and V_{s30} (m/s). PGA, PGV and PGD are the ground-motion parameters to be formulated. Fig. 3 shows the distribution of the data used to develop the predictive equations. For more visualization, the data are presented by frequency histograms (Fig. 4). As can observe from Fig. 4, the distributions of the predictor variables are not uniform. The derived models would most probably provide better predictions for the cases where the

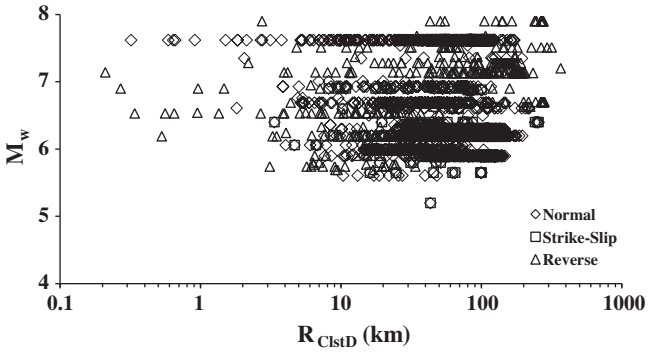

Fig. 3. Distribution of the data used in the model development, differentiated by fault type.

densities of the variables are higher. The ranges and statistics of different parameters involved in the modeling process are also given in Table 1. The amount of data used for the training process is an important issue as it guarantees the reliability of final models.

For the analysis, the available data sets are randomly divided into learning, validation and testing subsets. The learning data were taken for the training of the algorithm. The validation sets are used to specify the generalization capability of the models on data they did not train on (model selection). Thus, both of the learning and validation data are involved in the modeling process and were categorized into one group referred to as "training data" [1]. Finally, the testing data are employed to measure the performance of the models obtained by ANN/SA on the data that played no role in building the models. To obtain a consistent data division, several combinations of the training (learning and validation) and testing sets are considered. The selection is such that the maximum, minimum, mean and standard deviations of the parameters are consistent in the training and testing data sets [1]. Of the 2815 data sets, 2252 data vectors (80%) are taken for the training process (1971 sets for learning and 281 sets for validation). The remaining 563 data (20%) are used for the testing of the models.

Both the input and output variables are normalized in this study. After controlling several normalization methods [51], the following method is used to normalize the variables to a range of $[L,U]$:

$$X_n = ax + b \tag{10}$$

where,

$$a = \frac{U - L}{X_{\max} - X_{\min}} \tag{11}$$

$$b = U - aX_{\max} \tag{12}$$

in which $X_{\max}$ and $X_{\min}$ are the maximum and minimum values of the variable and X_n is the normalized value. In the present study, $L = 0.05$ and $U = 0.95$.

3.2. Performance measures

The best ANN/SA models are chosen on the basis of a multi-objective strategy as below [1]:

i. Selecting the simplest model, although this is not a predominant factor.
ii. Providing the best fitness value on the learning data sets.
iii. Providing the best fitness value on the validation data sets.

The first objective can be controlled by the user through the parameter settings (e.g., chromosome length). For the other objectives, the following objective function (OBJ) is constructed as a measure of how well the model predicted output agrees with the measured output. The selections of the best models are deduced by the minimization of following function [1]:

$$OBJ = \left(\frac{\mathrm{No.}_{\mathrm{Learning}} - \mathrm{No.}_{\mathrm{Validation}}}{\mathrm{No.}_{\mathrm{Training}}}\right)\frac{MAE_{\mathrm{Learning}}}{R^2_{\mathrm{Learning}}} + \frac{2\mathrm{No.}_{\mathrm{Validation}}}{\mathrm{No.}_{\mathrm{Training}}} \times \frac{MAE_{\mathrm{Validation}}}{R^2_{\mathrm{Validation}}} \tag{13}$$

where $\mathrm{No.}_{\mathrm{Learning}}$, $\mathrm{No.}_{\mathrm{Validation}}$ and $\mathrm{No.}_{\mathrm{Training}}$ are respectively the number of the learning, validation and training data. R and MAE are respectively correlation coefficient and mean absolute error given in the form of formulas as follows:

$$R = \frac{\sum_{i=1}^{n}(h_i - \bar{h}_i)(t_i - \bar{t}_i)}{\sqrt{\sum_{i=1}^{n}(h_i - \bar{h}_i)^2 \sum_{i=1}^{n}(t_i - \bar{t}_i)^2}} \tag{14}$$

$$MAE = \frac{\sum_{i=1}^{n}|h_i - t_i|}{n} \tag{15}$$

in which h_i and t_i are respectively the actual and predicted output values for the ith output, $\bar{h}_i$ and $\bar{t}_i$ are the average of the actual and predicted outputs, and n is the number of sample. The constructed objective function takes into account the changes of R and MAE together. High R values and low MAE values result in lowering OBJ and, consequently, indicate a more precise model. In addition, the above function considers the effects of different data divisions for the learning and validation data. Mean absolute percent error (MAPE) and mean squared error (MSE) are also calculated using the following equations:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left[\left|\frac{h_i - t_i}{h_i}\right|\right] \tag{16}$$

$$MSE = \frac{\sum_{i=1}^{n}(h_i - t_i)^2}{n} \tag{17}$$

3.3. Model development using the hybrid ANN/SA method

The available database is used for developing the ANN/SA prediction models relating PGA, PGV and PGD to F, M_w, R_{ClstD} (km), and V_{s30} (m/s). In addition to separate models, combined models are also established using PGA, PGV and PGD as the outputs. Several runs are conducted to come up with an efficient parameterization for ANN/SA. The parameters involved in the ANN/SA algorithm are changed for different runs. These parameters are selected based on some previously suggested values [28] and after a trial and error approach. As discussed above, choosing an appropriate acceptance constant (k) is important for training ANN/SA. The network is trained using both of the temperature cycling and linear cooling schedules. A value of $k = 1500$, for a cooling schedule using an initial temperature of 15 and a final temperature of 0.015, provided reasonable learning for the investigated problem. Three different levels (5, 10 and 20) are considered for the number of cycles before starting the optimization process. Relatively low values (20, 30 or 50) are selected for the number of iterations at each temperature for the temperature cycling schedule. The performance of an ANN/SA model mainly depends on the network architecture and parameter settings. According to a universal approximation theorem [32], a single hidden layer network is sufficient for the ANN-based methods (such as hybrid ANN/SA) to uniformly approximate any continuous and nonlinear function. Choosing the number of hidden layers, hidden nodes, learning rate, epochs, and activation function type plays an important role in the model construction. Hence, several networks with different settings for the mentioned

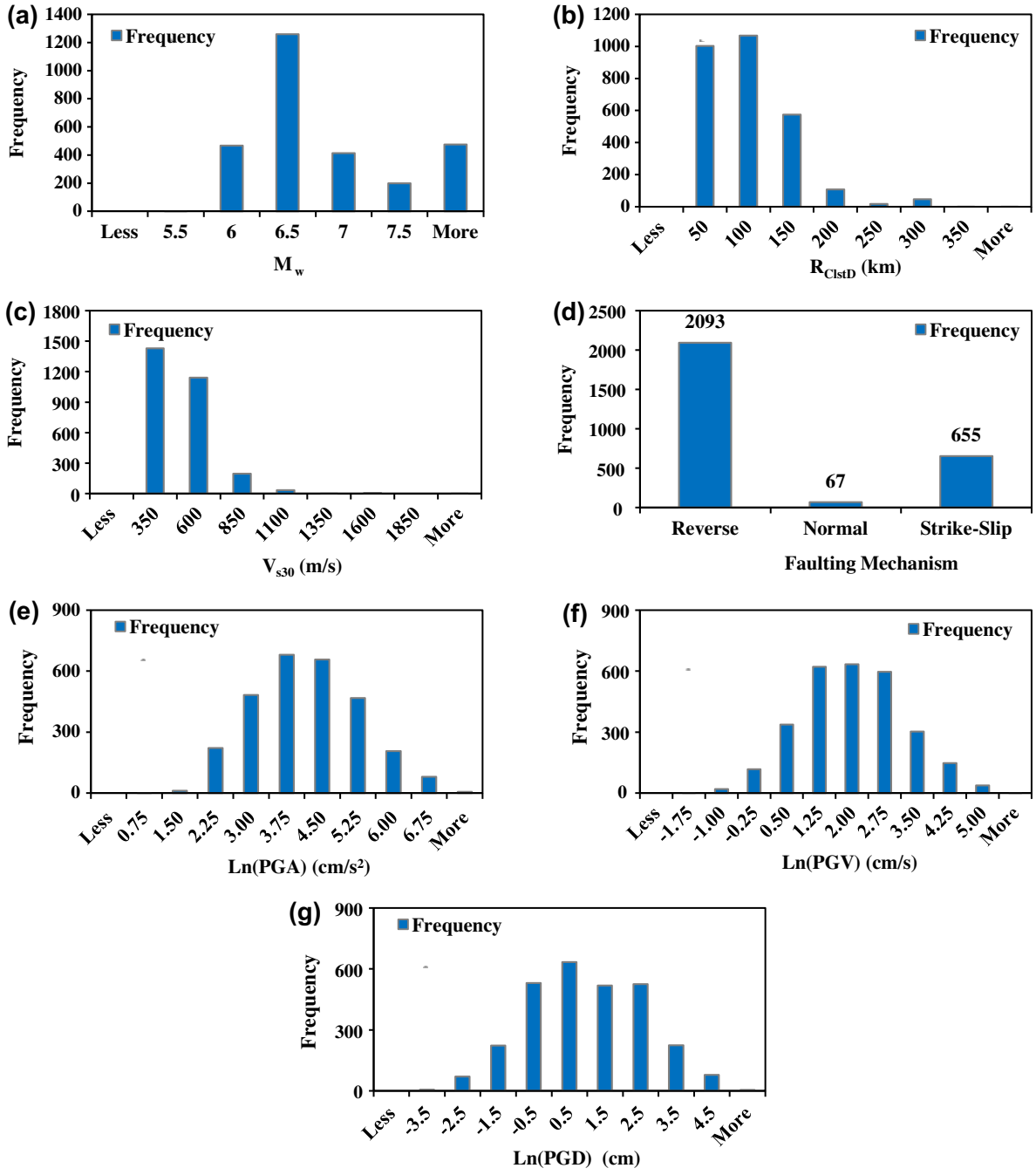

Fig. 4. Histograms of the variables used in the model development.

Table 1
Descriptive statistics of the variables used in the model development.

Parameter	M_w	R_{ClstD} (km)	V_{s30} (m/s)	F	PGA (cm/s^2)	PGV (cm/s)	PGD (cm)
Mean	6.55	73.80	387.19	1.49	85.06	10.08	5.28
Standard error	0.01	0.95	3.30	0.02	2.34	0.27	0.21
Median	6.30	63.79	345.42	1	43.13	5.01	1.48
Standard deviation	0.59	50.24	175.09	0.85	124.04	14.46	11.21
Sample variance	0.35	2524.41	30654.89	0.72	15385.32	209.02	125.67
Kurtosis	−0.71	3.20	11.99	−0.54	27.24	18.24	113.84
Skewness	0.81	1.39	2.28	1.19	4.19	3.59	7.91
Range	2.70	365.96	1899.78	2	1628.83	169.86	232.38
Minimum	5.20	0.07	116.35	1	1.14	0.10	0.01
Maximum	7.90	366.03	2016.13	3	1629.97	169.96	232.39

characters are trained to reach the optimal configurations with the desired precision. Conjugate-gradient and Levenberg–Marquardt algorithms are implemented for the training of the network. In all cases, the best results are obtained by the Levenberg–Marquardt method. Also, the transfer function between the input and hidden layer is log-sigmoid of form $1/(1+e^{-x})$. A linear transfer function (purelin) is adopted between the hidden layer and output layer.

The weights and biases are randomly assigned for each run. These assignments considerably change the performance of a newly trained network even all the previous parameter settings and the architecture are kept constant. This leads to extra difficulties in selection of optimal architecture and parameter settings. To overcome this difficulty, the weights and biases are frozen after the networks were well trained. Thereafter, the following function is used to convert the optimal ANN/SA models into mathematical equations relating the input parameters and the output parameter (h) [52]:

$$h = f_{HO}\left(bias_{hk} + \sum_{k=1}^{h} V_k f_{IH}\left(bias_{hk} + \sum_{i=1}^{m} w_{ik}x_i\right)\right) \tag{18}$$

where $bias_h$ = hidden layer bias; V_k = weight connection between neuron k of the hidden layer and the single output neuron; $bias_{hk}$ = bias at neuron k of the hidden layer ($k = 1,h$); w_{ik} = weight connection between the input variable ($i = 1,m$) and neuron k of the hidden layer; x_i = input parameter i; f_{HO} = transfer function between the hidden layer and output layer; and f_{IH} is the transfer function between the input and hidden layer. The ANN/SA algorithm is implemented using the Neural-Lab program (Version 3.1) [53].

3.3.1. ANN/SA-based ground-motion attenuation models

The best model architecture for the prediction of the ground-motion parameters is found to contain:

- Temperature cycling cooling schedule for the initialization phase.
- One invariant input layer, with 4 ($n = 4$) arguments (F, M_w,
- R_{ClstD}, V_{s30}) and a bias term;
- One invariant output layer with 3 nodes providing the value of Ln(PGA), Ln(PGV) and Ln(PGD).
- One hidden layer having 8 ($m = 8$) nodes.

The ANN/SA-based formulations of PGA (cm/s^2), PGV (cm/s) and PGD (cm) are as follows:

$$\begin{pmatrix} \ln(PGA) \\ \ln(PGV) \\ \ln(PGD) \end{pmatrix} = \begin{pmatrix} \frac{1}{0.1238}\left(bias_h - 0.034 + \sum_{k=1}^{8} \frac{V_k}{1+e^{-F_j}}\right) \\ \frac{1}{0.1206}\left(bias_h - 0.3299 + \sum_{k=1}^{8} \frac{V_k}{1+e^{-F_j}}\right) \\ \frac{1}{0.0909}\left(bias_h - 0.4545 + \sum_{k=1}^{8} \frac{V_k}{1+e^{-F_j}}\right) \end{pmatrix} \tag{19}$$

where,

$$F_j = F_n \times W_{1k} + M_{w,n} \times W_{2k} + \ln(R_{ClstD})_n \times W_{3k} + V_{s30,n} \times W_{4k} + bias_k \tag{20}$$

in which, F_n, $M_{w,n}$, $\ln(R_{ClstD})_n$, and $V_{s30,n}$ represent the inputs variables normalized using Eq. (10). k is the number of the hidden layer neurons. The input layer weights (W_k), input layer biases ($bias_k$), hidden layer weights (V_k), and hidden layer biases ($bias_h$) of the optimum ANN/SA model are presented in Tables 2 and 3. The model is trained for 364 epochs. Fig. 5 shows the network's MSE as a function of time for the normalized output data. It can be observed from this figure that MSE is remarkably minimized during the initialization phase. This implies that SA has offered very good initial values for the weights. Once the initialization phase is completed, the training process switches to the optimization phase. Since the network reaches an acceptable MSE value (0.0057) during the initialization phase, the optimization process attains a very low MSE equal to 0.0048 after 240 s. A comparison of the measured and predicted Ln(PGA), Ln(PGV) and Ln(PGD) values by ANN/SA is shown in Figs. 6–8, respectively.

3.4. Model development using single ANN

A new single ANN model is developed to highlight the significance of using SA for the neural network training. The initialization process for single ANN involves assigning initial random values to the weights of the network instead of assigning good starting values to the weights by means of SA. In order to conduct a fair comparison, the architecture of single ANN is considered to be similar to that of ANN/SA. Therefore, the single ANN prediction model had 4 neurons in its input layer (F, M_w, R_{ClstD}, V_{s30}) and 3 neurons in the output layer (Ln(PGA), Ln(PGV) and Ln(PGD)). The ANN architecture is built with one hidden layer with 8 neurons and trained for 728 epochs. The network is trained using the Levenberg–Marquardt method. The log-sigmoid and linear functions are the transfer functions used for the input-hidden and hidden-output layers, respectively. The training and testing data sets used for developing the single ANN model are the same as those considered for establishing the ANN/SA model. A script is written in MATLAB environment using Neural Network Toolbox 5.1 to implement ANN. Figs. 9–11 visualize a comparison between the measured Ln(PGA), Ln(PGV) and Ln(PGD) values and the values predicted by the single ANN model, respectively.

3.5. Model development using regression analysis

A multivariable least squares regression (MLSR) analysis is performed to have an idea about the predictive power of the best ANN/SA models. The method of LSR is extensively used in regression analysis primarily because of its interesting nature. Under certain assumptions, LSR has some attractive statistical properties that have made it as a member of the most powerful and popular methods of regression analysis. LSR minimizes the sum-of-squared residuals for each equation, accounting for any cross-equation restrictions on the parameters of the system. The MLSR models are trained using the total of the learning and validation data sets considered for developing the ANN/SA models. The testing sets used for evaluating the generalization capabilities of the MLSR models are the same as those employed for testing ANN/SA. Eviews software package [54] is used to perform the regression analysis. The MLSR-based formulations of PGA (cm/s^2), PGV (cm/s) and PGD (cm) in terms of F, M_w, $\ln(R_{ClstD})$, and V_{s30} are as given below:

$$\begin{pmatrix} \ln(PGA) \\ \ln(PGV) \\ \ln(PGD) \end{pmatrix} = a_1 F + a_2 M_w + a_3 \ln(R_{ClstD}) + a_4 V_{s30} + a_5 \tag{21}$$

where a denotes coefficient vector. The coefficient vectors for PGA, PGV and PGD are presented in Table 4. A comparison of the measured versus MLSR predicted Ln(PGA), Ln(PGV) and Ln(PGD) values is shown in Figs. 12–14, respectively.

4. Performance analysis and model validity

Different models are developed for the estimation of PGA, PGV and PGD upon a reliable database. Based on a rational hypothesis [55], if a model gives $|R| > 0.8$ and the error values (e.g., MAE and MAPE) are minimized, a strong correlation exists between the predicted and measured values. It can be observed from Figs. 6–8 that

Table 2
Weight and bias values between the input and hidden layer.

Weights	Number of hidden neurons (k)							
	1	2	3	4	5	6	7	8
W_{1k}	1.9448	21.8787	−0.2996	0.7318	−0.8796	−24.9341	1.2175	6.9636
W_{2k}	−5.6354	−269.5708	3.0580	15.7272	3.4759	−44.0730	4.8294	−3.9554
W_{3k}	−1.7975	71.9498	−12.9053	−0.1874	47.9821	−22.2205	−3.4144	−0.8005
W_{4k}	−0.6256	7782.5995	−0.4807	0.2956	−8.3088	−16.2937	−1.6121	−1.7217
$bias_k$	12.7706	−756.1658	9.0881	−2.8306	−38.8312	41.7184	0.5670	5.2045

Table 3
Weight and bias values between the hidden and output layer.

Weights		Number of hidden neurons (k)								
		1	2	3	4	5	6	7	8	$bias_h$
V_k	PGA	326.4047	−0.0239	0.6294	−0.3338	0.1040	0.0356	0.2412	−1.3943	−324.7352
	PGV	341.1122	−0.0574	0.4766	−0.2935	0.1264	0.0852	0.4895	−1.7468	−339.1730
	PGD	249.9431	−0.0598	0.2435	−0.1880	0.1000	0.1059	0.6541	−1.5400	−248.3047

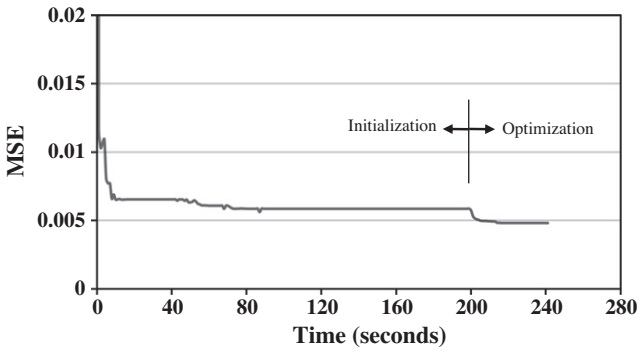

Fig. 5. Variations of the MSE values with time during the ANN/SA training process.

the ANN/SA models with high R and low MAE and MAPE values provide acceptable estimates of the target values. The performance of the models on the training and testing data indicates that they have both good predictive ability and generalization performance. As can be seen in Figs. 6–14, ANN/SA has provided notably better results than the single ANN and regression methods, in particular for the prediction of PGA. The results clearly indicate that using SA to adjust the network's weights improves the performance of single ANN.

Frank and Todeschini [56] state that the minimum ratio of the number of objects over the number of selected variables for model acceptability is 3. It is also suggested that considering a ratio equal to 5 would be safer. In the present study, this ratio is much higher and is equal to 2815/4 = 703.8.

Furthermore, the external validation of the ANN/SA models on the testing data sets is evaluated new criteria recommended by Golbraikh and Tropsha [57]. It is suggested that at least one slope of regression lines (k or k') through the origin should be close to 1. Also, the performance indexes of m and n should be lower than 0.1. Either the squared correlation coefficient (through the origin) between predicted and experimental values (Ro^2), or the coefficient between experimental and predicted values (Ro'^2) should be close to 1 [1,57]. The considered validation criteria and the relevant results obtained by the models are presented in Table 5. With the exception of the m and n criteria, the proposed models satisfy the other conditions. The model validity results indicate that the derived ANN/SA models are valid and can reliably be used to determine the principal ground-motion parameters. It is notable that the derived models mainly have a predictive capability within the data range used for their calibration.

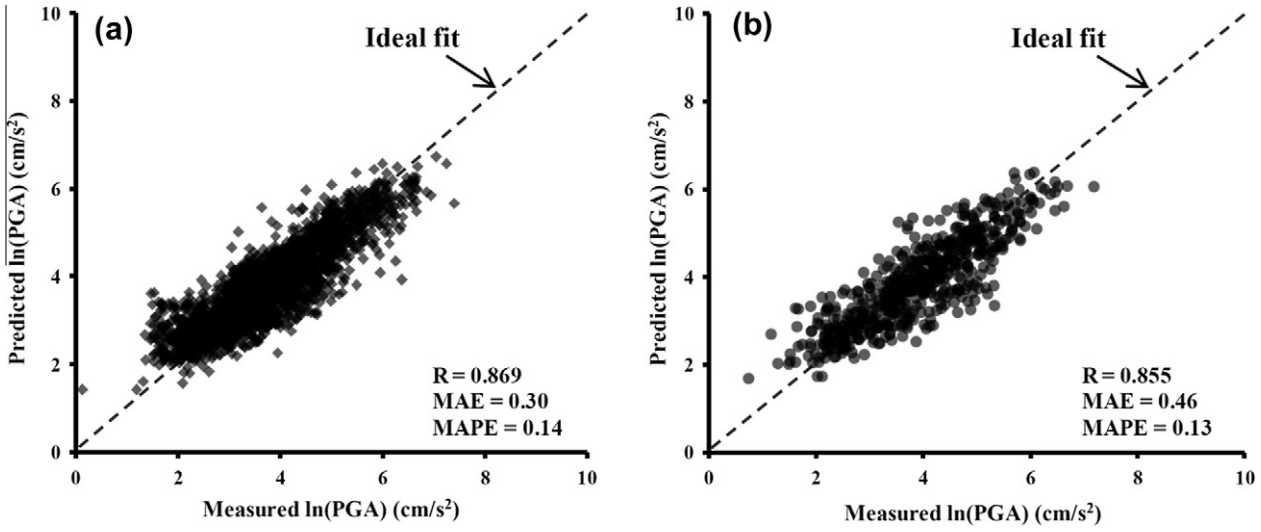

Fig. 6. Measured versus predicted Ln(PGA) values using the ANN/SA model: (a) training data, (b) testing data.

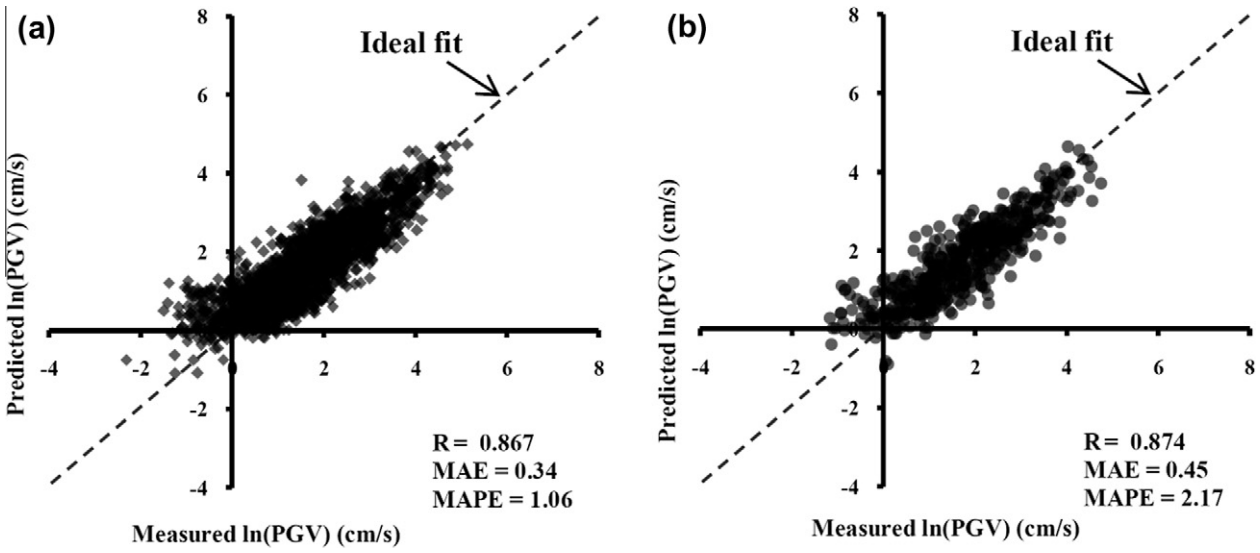

Fig. 7. Measured versus predicted Ln(PGV) values using the ANN/SA model: (a) training data, (b) testing data.

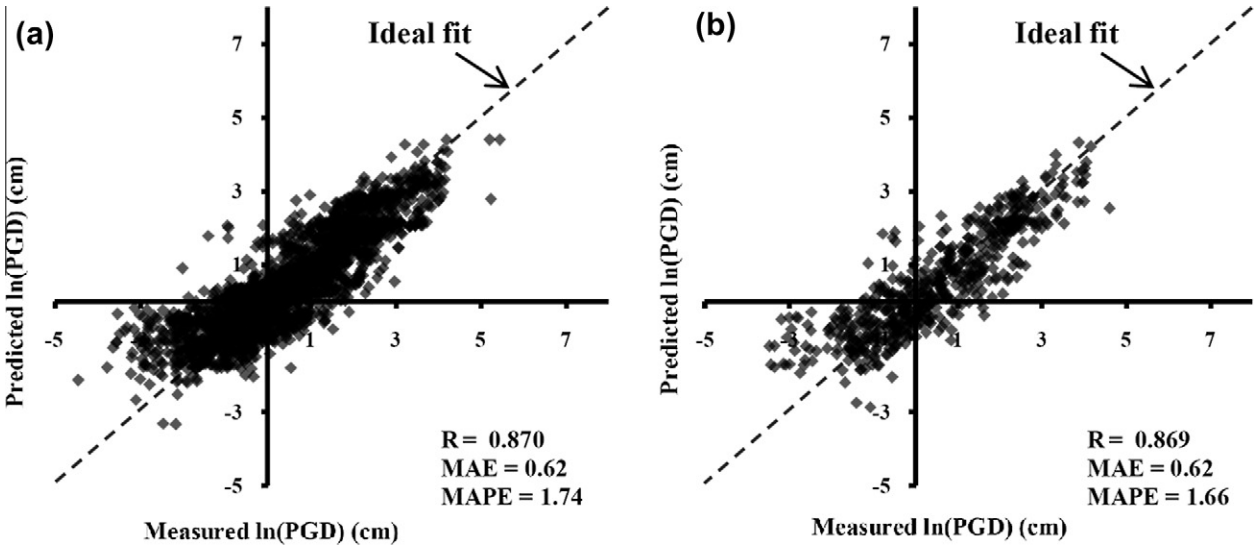

Fig. 8. Measured versus predicted Ln(PGD) values using the ANN/SA model: (a) training data, (b) testing data.

5. Comparison with the existing ground-motion prediction models

There are numerous prediction models for PGA, PGV and PGD [5,13]. A comparison of the ANN/SA attenuation relations with all of the existing ground-motion models is beyond the scope of this study. Instead, the results obtained herein are compared with those provided by the well-known models of Ambraseys et al. [7] and Smit et al. [29] for PGA. As a part of the PEER Next Generation Attenuation of Ground Motion (NGA) Project, Campbell and Bozorgnia [10] developed updated empirical models for PGA, PGV and PGD. The results of this recent research were included in the comparative study. The reason for choosing these models is that they belong to different generations. Therefore, they can be regarded as good benchmarks for the ANN/SA models.

An efficient way to observe the systematic mismatches between the predictions and observations is to plot the residuals. The residuals are defined as the ratios of the measured to predicted ground-motion parameters [9]. Figs. 15–17 respectively show the residuals as a function of distance for PGA, PGV and PGD for different earthquake magnitudes. In order to conduct a fair comparison, the performance of the models is evaluated on 563 data sets used for the testing of the ANN/SA models. The considered performance measures are MAPE and the values of mean (Mean) of the residuals. As can be observed from these figures, the ANN/SA attenuation models for PGA and PGV outperform the regression-based models of Ambraseys et al. [7], Smit et al. [29] and Campbell and Bozorgnia [10]. It should be noted that the models developed by Ambraseys et al. [7] and Campbell and Bozorgnia [10] provide the best Mean values for PGA and PGV, respectively. Considering the results for PGD, it can be seen in Fig. 17 that the ANN/SA model (MAPE = 0.89, Mean = 1.29) considerably outperforms the Campbell and Bozorgnia's NGA model (MAPE = 5.73, Mean = 0.33).

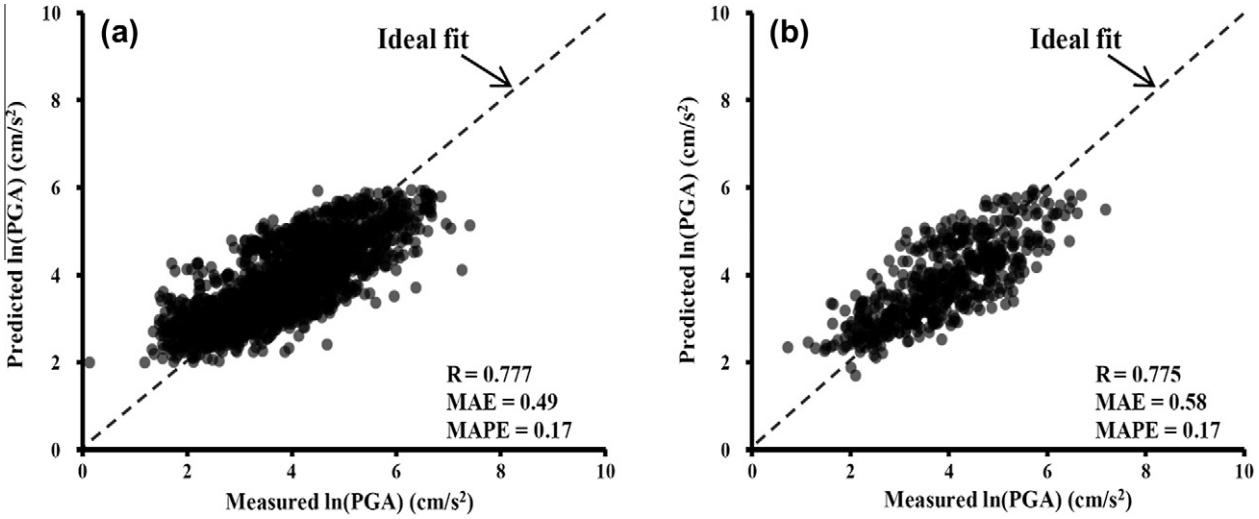

Fig. 9. Measured versus predicted Ln(PGA) values using the single ANN model: (a) training data, (b) testing data.

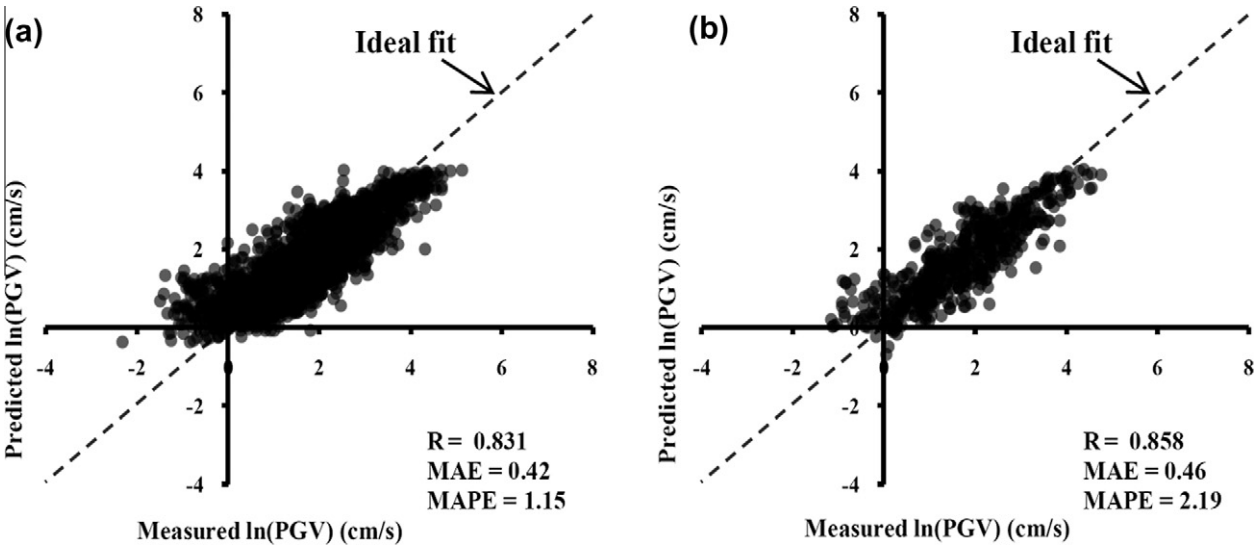

Fig. 10. Measured versus predicted Ln(PGV) values using the single ANN model: (a) training data, (b) testing data.

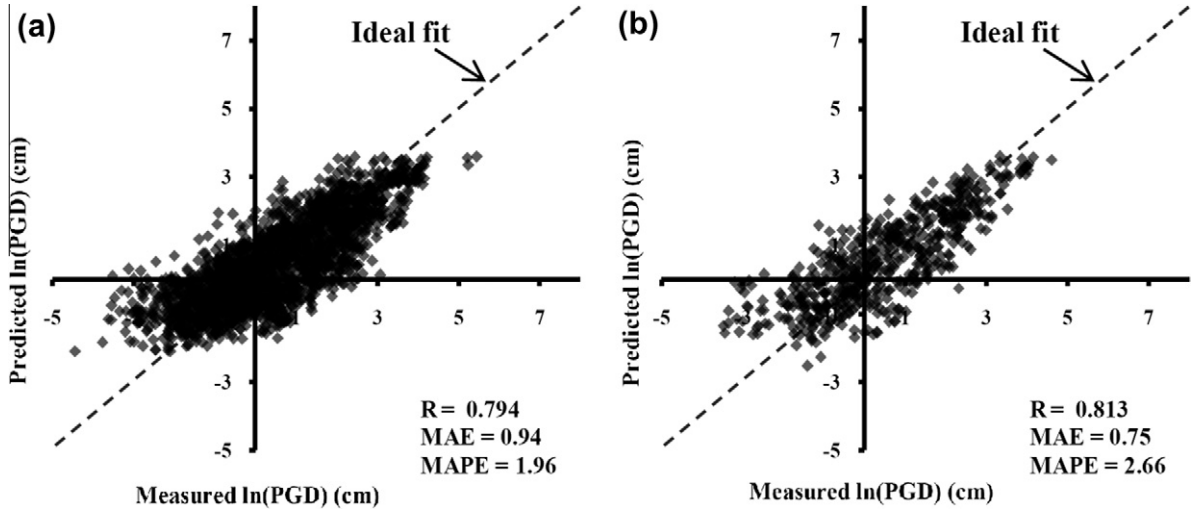

Fig. 11. Measured versus predicted Ln(PGD) values using the single ANN model: (a) training data, (b) testing data.

Table 4
The coefficient vectors derived for PGA, PGV and PGD.

Variable	α_1	α_2	α_3	α_4	α_5
PGA	−0.0859	0.5490	−0.9515	−0.0005	4.3679
PGV	−0.0471	1.0349	−0.8327	−0.0013	−1.2334
PGD	0.0842	1.9155	−0.7590	−0.0017	−8.4585

The PGA, PGV and PGD attenuation curves predicted by different models for the strike-slip faults are respectively illustrated in Figs. 18–20 for earthquakes of magnitude 6.5 and 7.5. As shown in Fig. 18, the PGA values predicted by the ANN/SA model for soil and rock sites are in good agreement with those provided by the Ambraseys et al. [7] and Campell and Bozorgnia [10] relationships. In this case, the predictions made by the model of Smit et al. [29] are far away from the others. For the rock sites and moment magnitude of 6.5 (Fig. 18b), the Ambraseys et al. [7] and Campell and Bozorgnia [10] attenuation curves closely follow the ANN/SA curve at distances below 40 km. A good consistency can be seen between the PGV values predicted by the ANN/SA and Campell and Bozrognia's NGA models (Fig. 19). The only notable difference lies in the PGV predictions for the soil sites for moment magnitude equal to 7.5 at distances over 30 km (Fig. 19c). The discrepancies in the PGA and PGV predictions can be assigned to different factors such as the differences in the amounts of data used for developing the relationships, distance definitions, soil categories, and fault type definitions.

As can be observed from Fig. 20, the PGD predictions made by the ANN/SA model are distant from those obtained by the Campell and Bozrognia's NGA model. The NGA model predicts greater PGD than the ANN/SA relationship. As previously shown in Fig. 17, the performance of the ANN/SA prediction model for PGD is significantly better than the NGA model. Therefore, in this case, the results obtained by the ANN/SA model can be regarded more reliable than those provided by the NGA model.

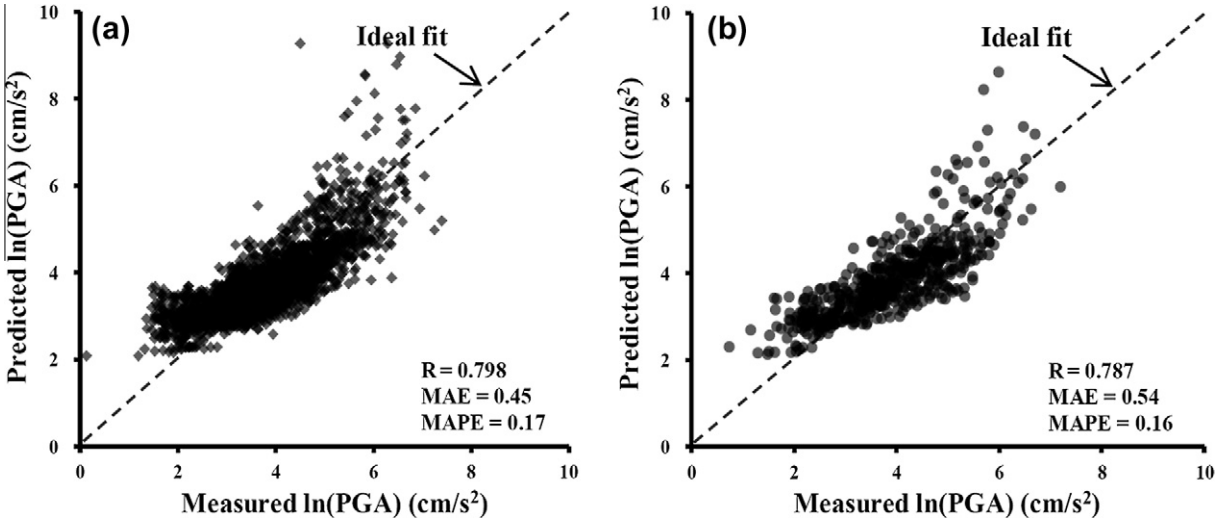

Fig. 12. Measured versus predicted Ln(PGA) values using the MLSR model: (a) training data, (b) testing data.

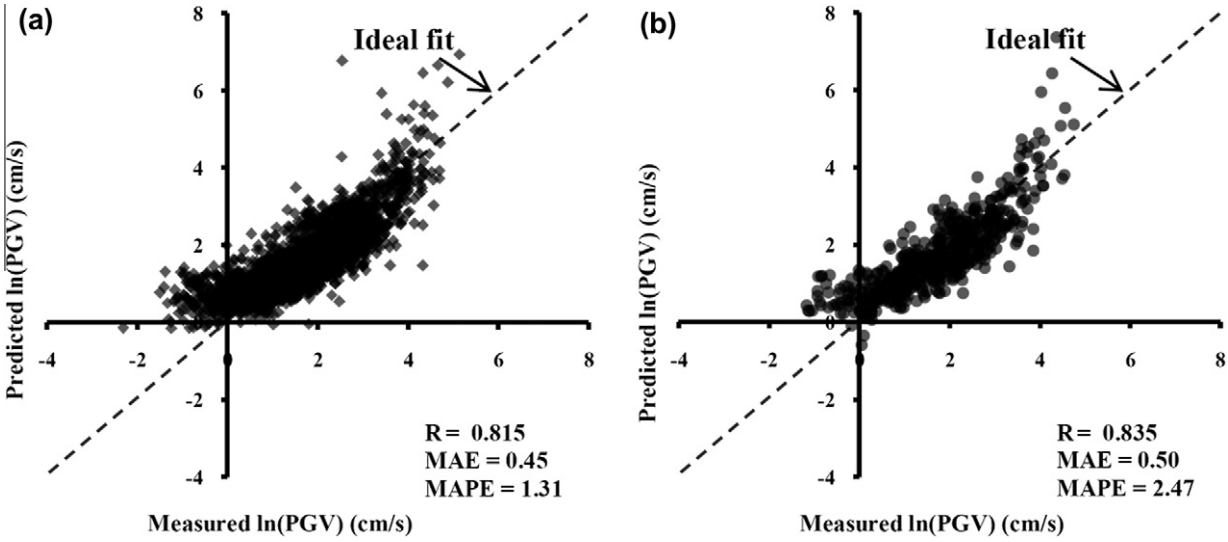

Fig. 13. Measured versus predicted Ln(PGV) values using the MLSR model: (a) training data, (b) testing data.

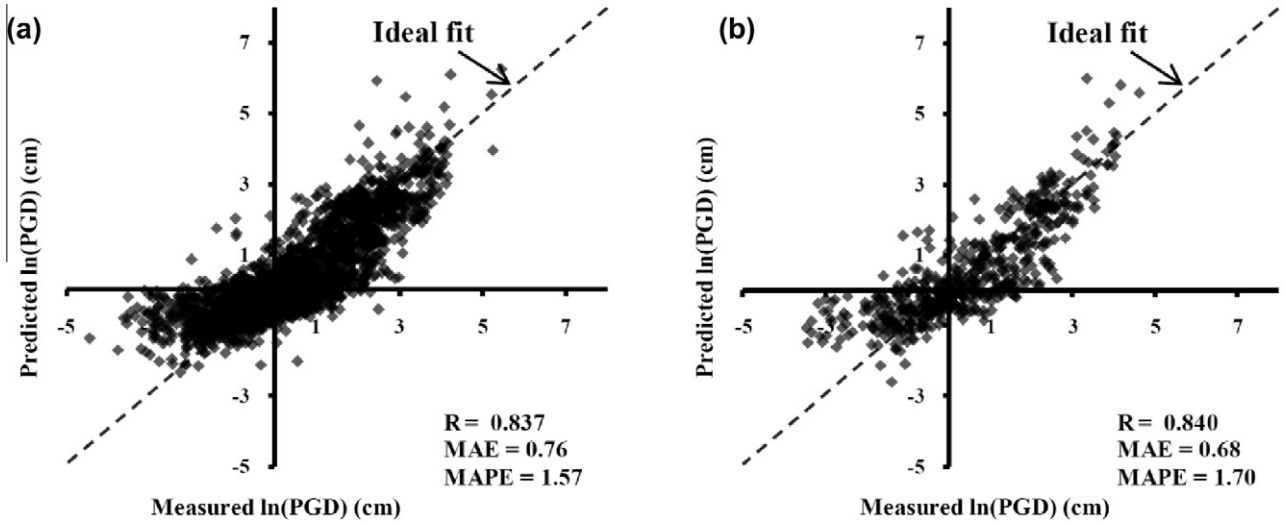

Fig. 14. Measured versus predicted Ln(PGD) values using the MLSR model: (a) training data, (b) testing data.

Table 5
Statistical parameters of the ANN/SA models for the external validation.

Item	Formula	Condition	PGA	PGV	PGD
1	R	$0.8 < R$	0.855	0.874	0.870
2	$k = \frac{\sum_{i=1}^{n}(h_i \times t_i)}{h_i^2}$	$0.85 < k < 1.15$	1.000	0.991	1.005
3	$k' = \frac{\sum_{i=1}^{n}(h_i \times t_i)}{t_i^2}$	$0.85 < k' < 1.15$	0.979	0.931	0.773
4	$m = \frac{R^2 - Ro^2}{R^2}$	$\lvert m\rvert < 0.1$	−0.368	−0.309	−0.322
5	$n = \frac{R^2 - Ro'^2}{R^2}$	$\lvert n\rvert < 0.1$	−0.358	−0.286	−0.245
where	$Ro^2 = 1 - \frac{\sum_{i=1}^{n}(t_i - h_i^o)^2}{\sum_{i=1}^{n}(t_i - \bar{t}_i)^2}$ and $h_i^o = k \times t_i$		1.000	1.000	1.000
	$Ro'^2 = 1 - \frac{\sum_{i=1}^{n}(h_i - t_i^o)^2}{\sum_{i=1}^{n}(h_i - \bar{h}_i)^2}$ and $t_i^o = k' \times h_i$		0.993	0.982	0.942

Note that the NGA models include several predictor (independent) variables such as the moment magnitude, one or more of the fault distance measures, indicator variables for style of faulting, hanging-wall parameters, shear-wave velocity, and one or more of the sediment depth parameters. On the other hand, the ANN/SA equations are developed using only four predictor variables (F, M_w, R_{ClstD}, and V_{s30}) and, therefore, can readily be used for routine design practice. As more data become available, including those for other main shocks or up-to-date NGA strong-motion data, the ANN/SA models can be improved to make more accurate predictions for a wider range.

Most of the existing ground-motion prediction models, such as those included in the comparative study, are derived by performing the regression analysis. These models assume the structure of the model in advance, which may be suboptimal. In most cases, the best regression models are obtained after controlling a few number of equations. Thus, they cannot efficiently consider the interactions between the predictor and output variables [1]. One of the major advantages of the ANN-based approaches over the traditional regression analysis is their ability to generate prediction models without any need to assume prior forms of the existing relationships. They directly learn from raw experimental (or field) data presented to them. It is notable that the expert systems, such as ANN/SA, are introduced into the design processes for better handling of information in pre-design phases [58]. In the first steps of design, it is idealistic to have some initial estimations of the outcome before conducting any extensive laboratory or field work. The ANN/SA system used herein is based on the data alone to determine the structure and parameters of the models. Hence, the results of the ANN/SA-based analyses are suggested to be treated as a complement to deterministic techniques. In order to develop a sophisticated prediction tool, ANN/SA can be combined with advanced deterministic geomechanical models. Assuming the deterministic model captures the key physical mechanisms, it needs appropriate initial conditions and carefully calibrated parameters to provide precise predictions. An efficient idea would be to calibrate the geomechanical parameters via ANN/SA which takes into account historic data sets as well as the laboratory or field test results.

6. Parametric analysis

For further verification of the ANN/SA-based prediction models, a parametric analysis is performed in this study. The parametric analysis investigates the response of the predicted ground-motion parameters to a set of seismic variables. The robustness of the design equations is determined by examining how well the predicted PGA, PGV and PGD values agree with the underlying physical behavior of ground-motions [19]. The PGA, PGV and PGD attenuation curves predicted by the ANN/SA models are shown in Fig. 21 for soil and rock sites for earthquakes of magnitude 6.5 and 7.5. As it is seen, PGA, PGV and PGD continuously increase due to

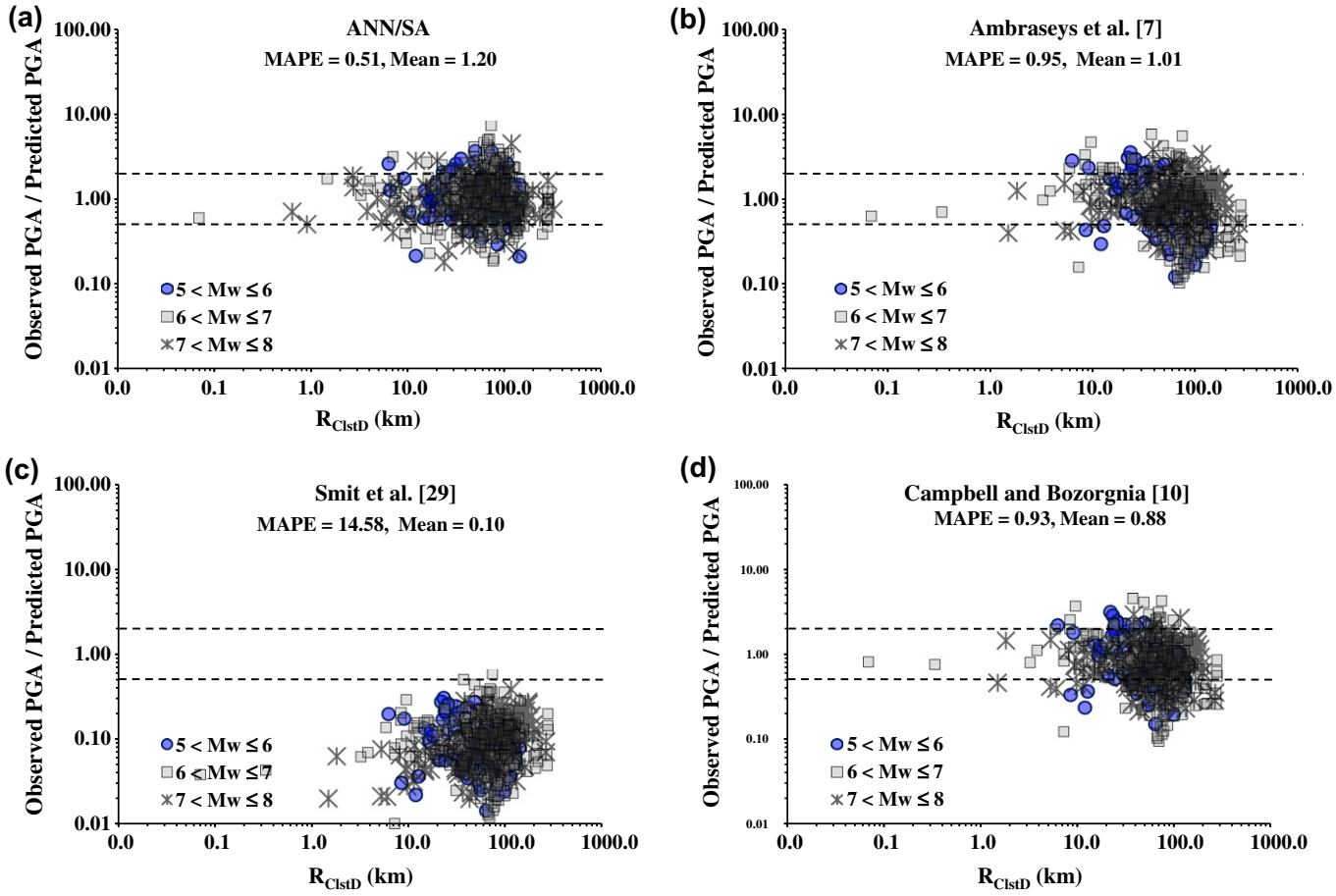

Fig. 15. The residual versus distance plots for PGA.

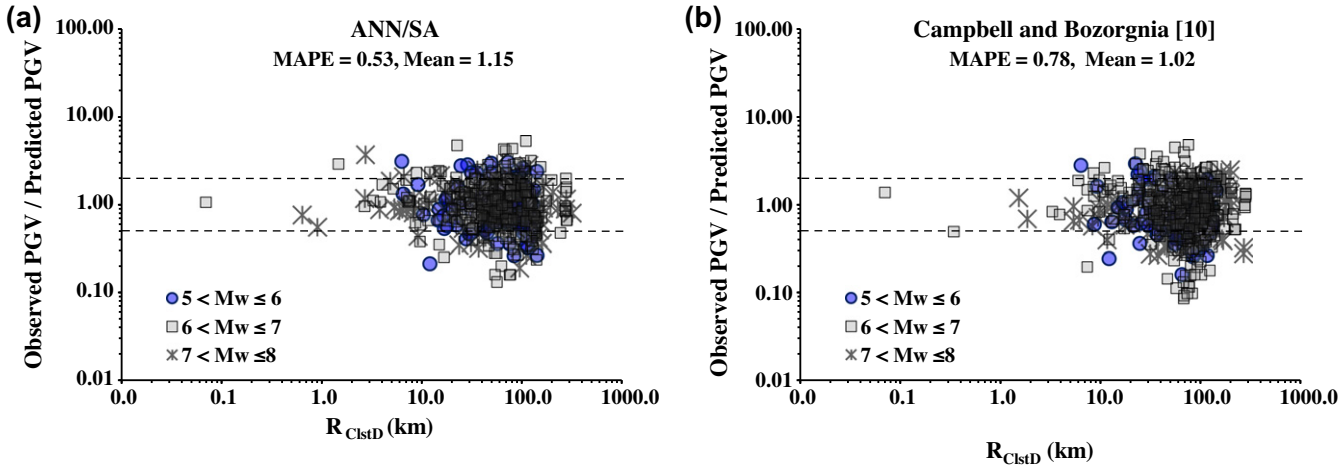

Fig. 16. The residual versus distance plots for PGV.

increasing M_w. Also, attenuation of PGA, PGV and PGD with distance is perfectly captured by the ANN/SA models. Soft soils amplify the ground-motions. As can be observed from these figures, the attenuation curves for the soil sites lie below those for the rock sites. The results of the parametric analysis are in close agreement with the theory of seismic wave attenuation. The results confirm that the proposed design equations are capable of capturing the important characteristics of peak ground-motion attenuation such as magnitude scaling, attenuation with distance and site amplification effects.

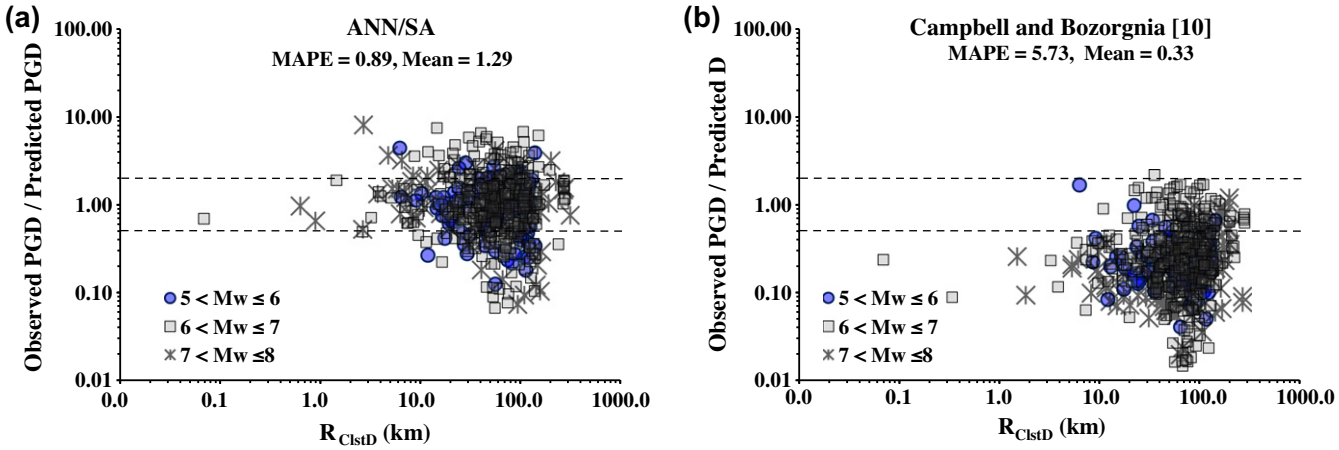

Fig. 17. The residual versus distance plots for PGD.

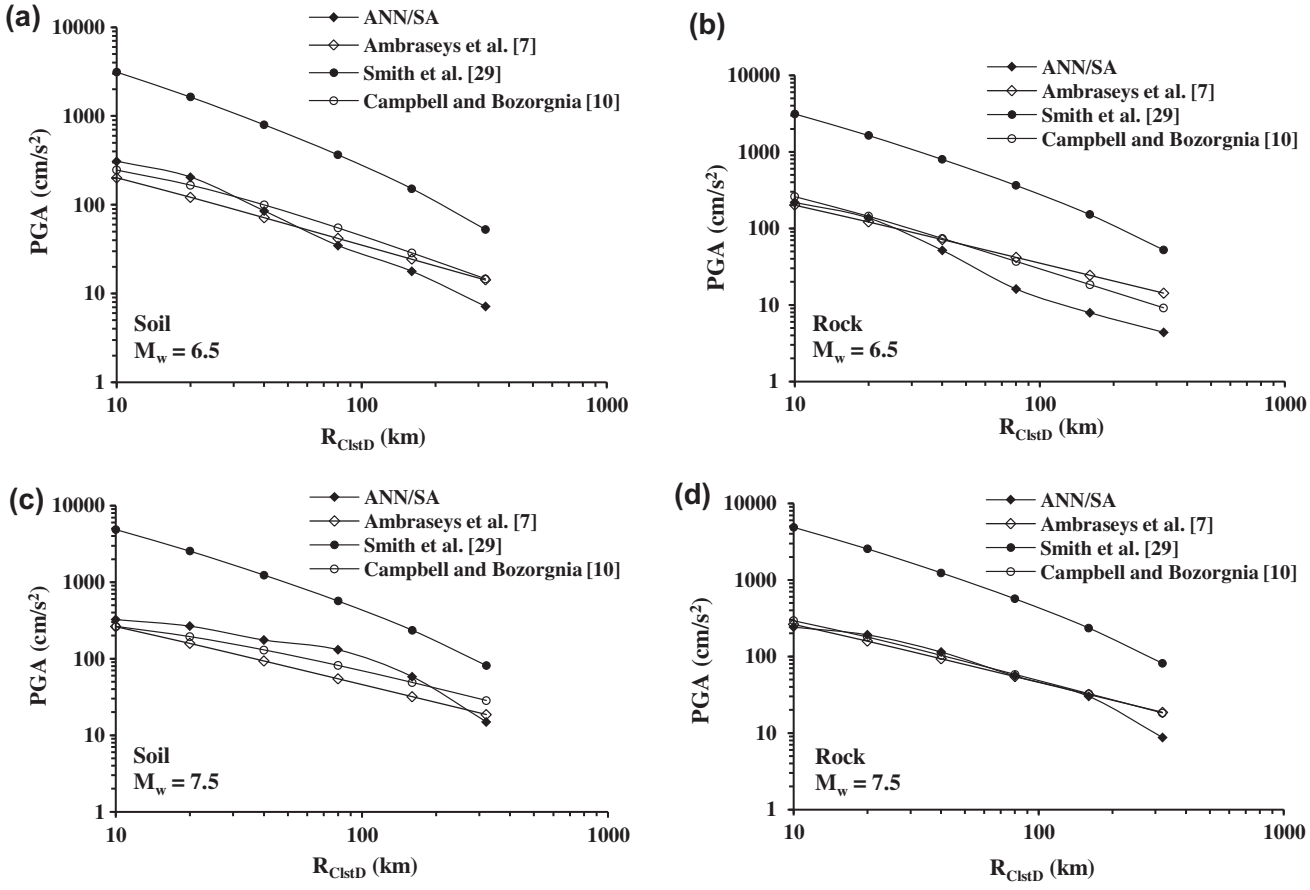

Fig. 18. The PGA-distance relations for soil and rock sites for different magnitudes.

7. Sensitivity analysis

The contribution of each input parameter in the ANN/SA models is evaluated through a sensitivity analysis. To achieve this, relative importance values of the predictor variables are calculated using Garson's algorithm [59]. A summary of the Garson's protocol for determining the relative importance values is shown in Fig. 22 [33]. Following this protocol, the input-hidden and hidden-output weights of the trained ANN/SA model are partitioned and the absolute values of the weights are taken to calculate the relative importance values.

The relative importance values of the input parameters are presented in Fig. 23. As it is seen, M_w and R_{ClstD} exert dominant influence on the variations of the ground-motion parameters. This

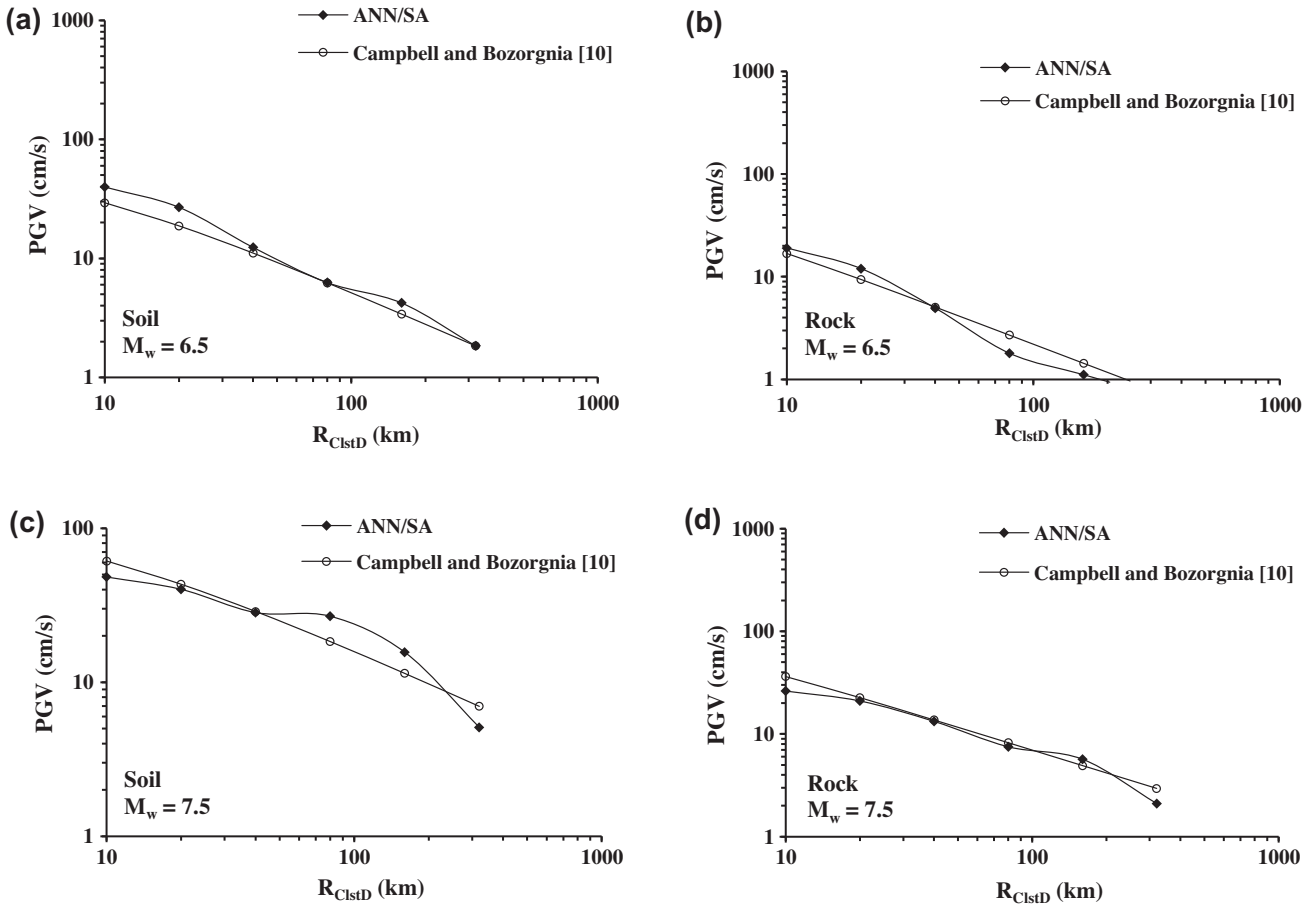

Fig. 19. The PGV-distance relations for soil and rock sites for different magnitudes.

is an expected case and goes arm in arm with the ground-motions behavior. Note that the fault type effect is significant in near-filed area (i.e., distances lower than 15 km).

8. Conclusions

In this research, a new hybrid method combining ANN and SA, referred to as ANN/SA, is utilized to derive prediction models for the time-domain ground-motion parameters. The proposed attenuation relationships are developed based on an extensive database containing thousands of records. The main conclusions of this paper can be drawn as follows:

i. The developed ANN/SA relationships provide reliable estimations of the PGA, PGV and PGD values. The models efficiently satisfy most of the conditions considered for their external validation.
ii. The proposed equations simultaneously take into account the role of several important factors (F, M_w, R_{ClstD}, and V_{s30}) representing the behavior of the strong ground-motion parameters.
iii. The tractable ANN/SA-based design equations provide analysis tools accessible to practicing engineers. The calculation procedure outlined in Appendix can readily be performed using a spreadsheet or hand calculations to give predictions of the PGA, PGV and PGD values.
iv. The proposed attenuation models can be used for practical pre-planning and design purposes since they were developed upon on a comprehensive database with wide range properties. The models are mostly valid for use in the western United States and in other similar tectonically active regions of shallow crusting faulting worldwide.
v. Further verification is done by comparing the proposed prediction models with the single ANN and MLSR solutions and the models developed by Ambraseys et al. [7], Smit et al. [29] and Campbell and Bozorgnia [10]. The ANN/SA models provide better prediction performance than the benchmark models. The proposed PGD prediction model produces considerably better outcomes than the Campbell and Bozorgnia's NGA model.
vi. Contrary to other optimization methods, SA is a no greedy optimization approach. Thus, it does not fall easily into local minima [28,44]. The results clearly indicate that introducing the SA strategy into the ANN modeling process results in a significant minimization of the error during the initialization phase.
vii. The performance of SA is strongly dependent on the cooling process [28]. The results show that using the temperature cycling cooling schedule for implementing SA leads to a more robust training in comparison with the traditional linear cooling schedule.
viii. The ANN/SA models capture the important characteristics of peak ground-motion attenuation including magnitude scaling, attenuation with distance and site amplification. The parametric study results are broadly consistent with the theory of seismic wave attenuation.

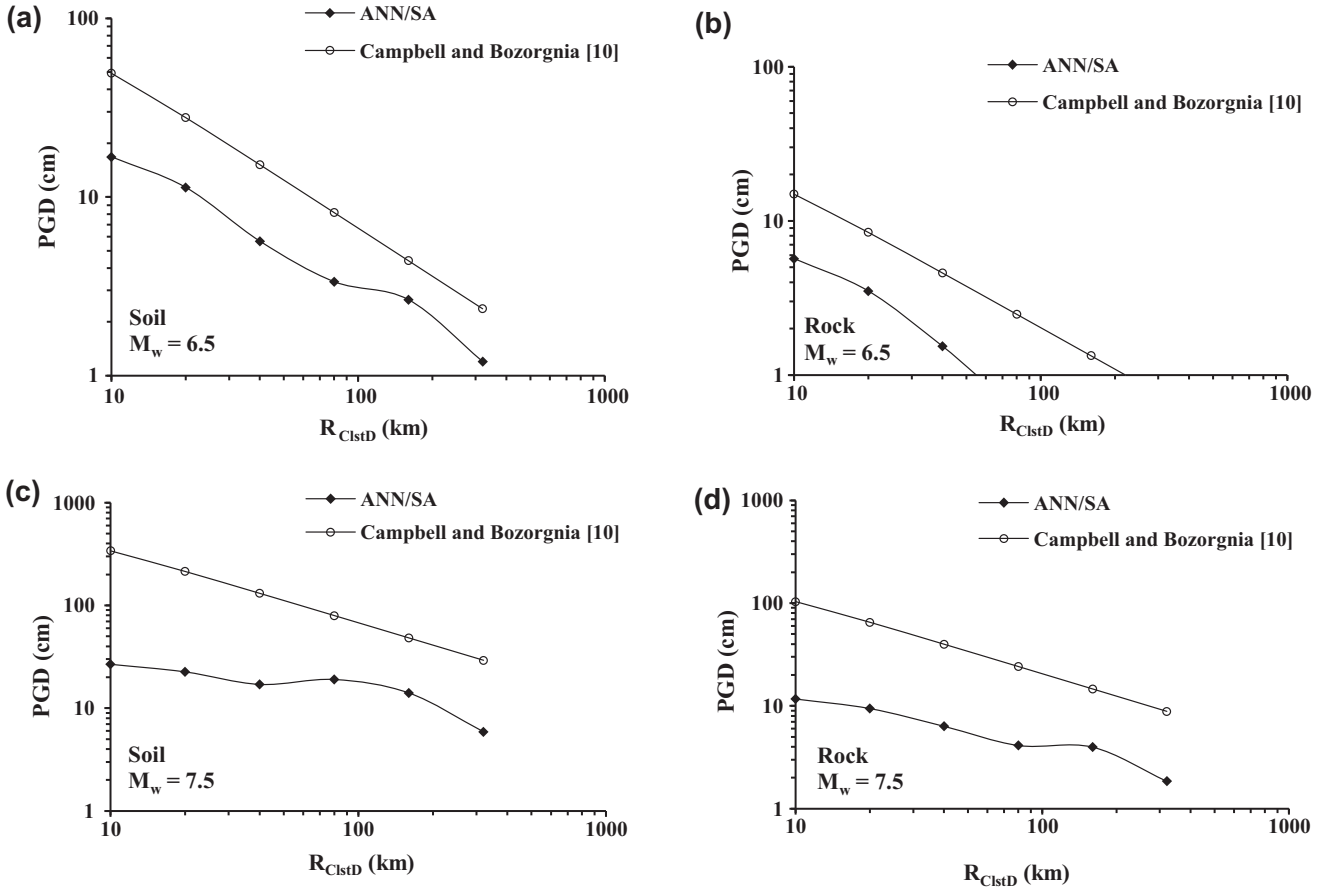

Fig. 20. The PGD-distance relations for soil and rock sites for different magnitudes.

ix. The sensitivity analysis results expectedly indicate that the most important parameters governing the ground-motion parameters behaviors are the earthquake magnitude and source to site distance.

x. A limitation of the conventional constitutive modeling methods is that they rely on assuming the structure of the model in advance. On the other hand, a distinctive feature of the ANN/SA technique is that it only uses the experimental data to specify the model structure.

Appendix A. Design example

An illustrative design example is provided to further explain the implementation of the ANN/SA ground-motion prediction equations. For this aim, one of the samples used for the testing of the models is taken. The M_w, R_{ClstD}, and V_{s30} values for this sample are, respectively, equal to 6.69, 5.19 km, and 370.52 m/s. The faulting mechanism is reverse (F = 1). Ln(PGA), Ln(PGV) and Ln(PGD) are required. The calculation procedure can be divided into three sections: (1) normalization of the input data, (2) calculation of the hidden layers, and (3) prediction of Ln(PGA), Ln(PGV) and Ln(PGD). The calculation procedure is outlined in the following steps:

Step 1: Normalization of the input data (F, M_w, R_{ClstD}, V_{s30}) to lie in a range from 0.05 to 0.95 and calculation of the input neurons (F_n, $M_{w,n}$, $R_{ClstD,n}$, $V_{s30,n}$) for each input data vector using Eqs. (10)–(12). The input neurons are calculated as: For F: the maximum and minimum values of the variable are 1 and 3, thus:

$$F_n = \left(\frac{0.95-0.05}{3-1}\right)F + \left(0.95 - \frac{0.95-0.05}{3-1}\times 3\right) = 0.050,$$

For M_w: the maximum and minimum values of the variable are 7.9 and 5.2, thus:

$$M_{w,n} = \left(\frac{0.95-0.05}{7.9-5.2}\right)M_w + \left(0.95 - \frac{0.95-0.05}{7.9-5.2}\times 7.9\right) = 0.547,$$

Similarly, $\mathrm{Ln}(R_{ClstD})_n = 0.503$ and $V_{s30,n} = 0.170$.

Step 2: Calculation of the hidden layer. The input value of each neuron in the hidden layer is determined for six neurons using the input layer weights and biases shown in Table 2. Given the information provided, the input values of the neuron ($F_1,\ldots,F_8$) are calculated using Eq. (20): $F_1 = 1.9448 \times 0.05 - 5.6354 \times 0.547 - 1.7975 \times 0.503 - 0.6256 \times 0.17 + 12.7706 = 8.775$ Similarly, $F_2 = 459.975$, $F_3 = 4.174$, $F_4 = 5.760$, $F_5 = -14.265$, $F_6 = 2.429$, $F_7 = 1.276$, and $F_8 = 2.695$.

Step 3: Prediction of Ln(PGA), Ln(PGV) and Ln(PGD). The input value of each output neuron is calculated using an activation function (log-sigmoid function). The calculated values are multiplied by the hidden layer connection weights (Table 3) and the summation is obtained:

$A_{PGA} = 326.4047f\ (F_1) - 0.0239f\ (F_2) + 0.6294f\ (F_3) - 0.3338f\ (F_4) + 0.1040f\ (F_5) + 0.0356f\ (F_6) + 0.2412f\ (F_7) - 1.3943f\ (F_8) - 324.7352 = 0.798.$

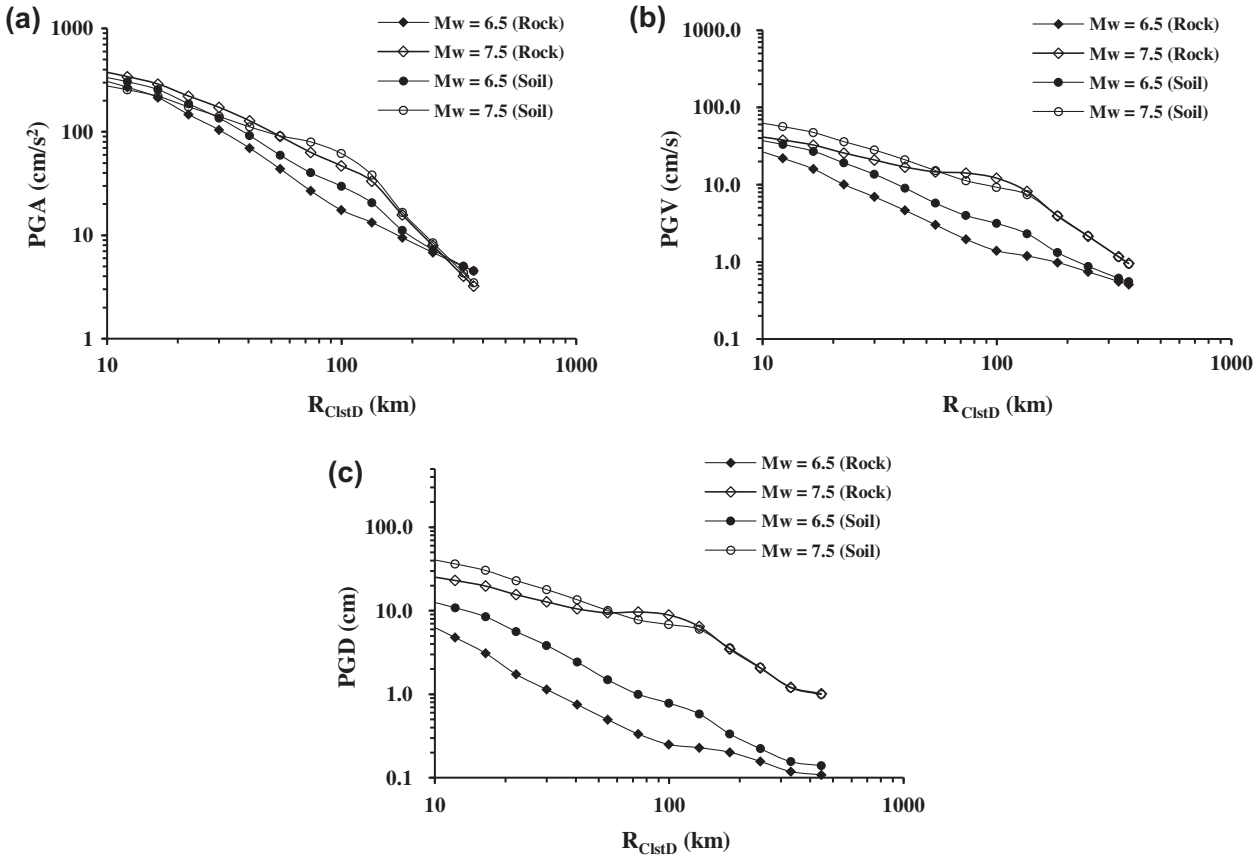

Fig. 21. Parametric analysis of the peak ground-motion parameters in the ANN/SA attenuation models: (a) PGA, (b) PGV, and (c) PGD.

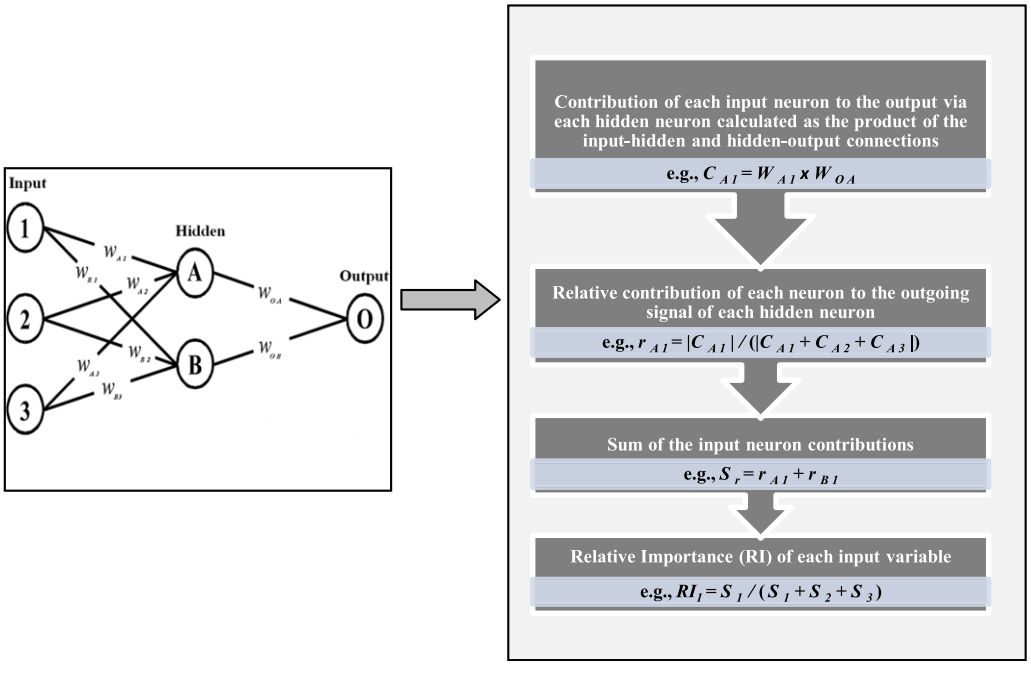

Fig. 22. Determining the relative importance of each input variable using the Garson's algorithm [33].

Similarly,

$A_{PGV} = 0.831$ and $A_{PGD} = 0.759$.

where $f(x)$ is the a log-sigmoid function of form $1/(1+e^{-x})$. Using Eq. (19), the values of Ln(PGA), Ln(PGV) and Ln(PGD) are calculated as follows:

$$\mathrm{Ln(PGA)} = \frac{1}{0.1238}(0.798 - 0.034) = 6.17 \text{ cm/s}^2$$

$$\mathrm{Ln(PGV)} = \frac{1}{0.1206}(0.831 - 0.3299) = 4.15 \text{ cm/s}$$

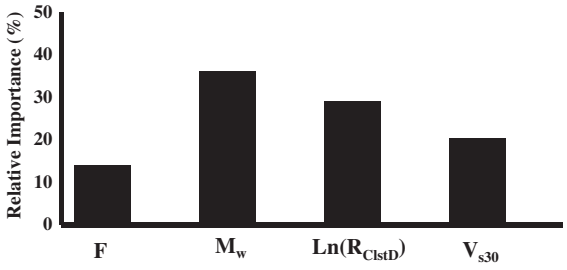

Fig. 23. Contributions of the predictor variables in the ANN/SA attenuation models.

$$\mathrm{Ln(PGD)} = \frac{1}{0.0909}(0.759 - 0.4545) = 3.35 \text{ cm}$$

In this example, the results are in good agreement with the measured values ($\mathrm{Ln(PGA)} = 6.45$ cm/s^2, $\mathrm{Ln(PGV)} = 4.55$ cm/s, $\mathrm{Ln(PGD)} = 3.51$ cm). The predicted Ln(PGA), Ln(PGV) and Ln(PGD) values are respectively 4.45%, 6.26% and 2.54% lower than the measured values.

References

[1] Gandomi AH, Alavi AH, Mousavi M, Tabatabaei SM. A hybrid computational approach to derive new ground-motion attenuation models. Eng Appl Artif Intell 2011;24(4):717–32.

[2] Luco N, Cornell CA. Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. Earthq Spectra 2007;23:357–92.

[3] Gullu H, Ercelebi E. A neural network approach for attenuation relationships: an application using strong-ground-motion data from Turkey. Eng Geol 2007;93:65–81.

[4] Krishnan S, Chen J, Komatitsch D, Tromp J. Case studies of damage to tall steel moment-frame buildings in Southern California during large San Andreas earthquakes. Bull Seismol Soc America 2006;96:1523–37.

[5] Douglas J. Earthquake ground motion estimation using strong-motion records: a review of equations for the estimation of peak ground acceleration and response spectral ordinates. Earth-Sci Rev 2003;61:43–104.

[6] Kramer SL. Geotechnical earthquake engineering. New Jersey: Prentice-Hall; 1996.

[7] Ambraseys NN, Simpson KA, Bommer JJ. Prediction of horizontal response spectra in Europe. Earthq Eng Struct Dyn 1996;25:371–400.

[8] Boore DM, Joyner WB, Fumal TE. Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes: A summary of recent work. Seismol Res Lett KNZ 1997:128–53.

[9] Boore DM, Atkinson GM. Boore–Atkinson NGA ground motion relations for the geometric mean horizontal component of peak and spectral ground motion parameters. PEER Report 2007/01, Pacific Engineering Research Center, University of California, Berkeley, CA, USA, 2007.

[10] Campbell KW, Bozorgnia Y. Campbell–Bozorgnia NGA ground motion relations for the geometric mean horizontal component of peak and spectral ground motion parameters. PEER Report 2007/02, Pacific Engineering Research Center, University of California, Berkeley, CA, USA, 2007.

[11] Fajfar P, Peruš I. A non-parametric approach to attenuation relations. J Earthq Eng 1997;1(2):319–40.

[12] Perus I, Fajfar P. Ground-motion prediction by a non-parametric approach. Earthq Eng Struct Dyn 2010;39(12):1395–416.

[13] Douglas J. Ground motion estimation equations 1964–2003. Research Report Number: 04-001-SM, Imperial College, UK, 2004.

[14] Cevik A, Cabalar AF. Modelling damping ratio and shear modulus of sand–mica mixtures using genetic programming. Exp Syst Appl 2009;36(4):7749–57.

[15] Kerh T, Chu D. Neural networks approach and microtremor measurements in estimating peak ground acceleration due to strong motion. Adv Eng Softw 2002;33:733–42.

[16] Chu D, Kerh T, Wu CH. Analysis of strong motion information at checking stations of Kaohsiung area by backpropagation neural networks. J Civil Eng Technol 2003;7:45–62.

[17] Kerh T, Ting SB. Neural network estimation of ground peak acceleration at stations along Taiwan high-speed rail system. Eng Appl Artif Intell 2005;8:857–66.

[18] Gunaydin K, Gunaydin A. Peak ground acceleration prediction by artificial neural networks for northwestern Turkey. Math Problems Eng 2008. ID 919420. doi:10.1155/2008/919420.

[19] Ahmad I, El Naggar MH, Naeem Khan A. Neural network based attenuation of strong motion peaks in Europe. J Earthq Eng 2008;12(5):663–80.

[20] Hamm L, Brorsen BW, Hagan MT. Comparison of stochastic global optimization methods to estimate neural network weights. Neural Process Lett 2007;26:145–58.

[21] Sexton RS, Dorsey RE, Johnson JD. Toward a global optimum for neural networks: a comparison of the genetic algorithm and backpropagation. Decision Support Syst 1998;22(2):171–86.

[22] Matilla-Garcia M, Arguello C. A hybrid approach based on neural networks and genetic algorithms to study the profitability in the Spanish stock market. Appl Econ Lett 2005;12:303–8.

[23] Sexton RS, Alidaee B, Dorsey RE, Johnson JD. Global optimization for artificial neural networks: a tabu search application. Euro J Oper Res 1998;106(2-3):570–84.

[24] Binner JM, Graham K, Gazely A. Co-evolving neural networks with evolutionary strategies. A new application to Divisia money. Adv Economet 2004;19:127–43.

[25] Porto VW, Fogel DB, Fogel LJ. Alternative neural network training models. IEEE Expert 1995:16–22.

[26] Sexton RS, Dorsey RE, Johnson JD. Beyond backpropagation: using simulated annealing for training neural networks. J End User Comput 1999;11:3–10.

[27] Ludermir TB. Neural networks for odor recognition in artificial noses. Proceedings of the International Joint Conference on Neural Networks, Portland, Oregon 2003, p. 143–148.

[28] Ledesma S, Torres M, Hernandez D, Avina G, Garcia G. Temperature cycling on simulated annealing for neural network learning. In: Proceedings of the MICAI 2007, LNAI 4827, 2007, p. 161–171.

[29] Smit P, Arzoumanian V, Javakhishvili Z, Arefiev S, Mayer-Rosa D, Balassanian S, Chelidze T. The digital accelerograph network in the Caucasus. In: Proceedings of the 2nd International Conference on Earthquake Hazard and Seismic Risk Reduction-Advances in Natural and Technological Hazards Research, Kluwer Academic Publishers, Yerevan, Armenia, 2000.

[30] Power MB, Chiou NA, Abrahamson N, Roblee C. The next generation of ground motion attenuation models, NGA project: an overview. In: Proceeding of the 8th National Conference on Earthquake Engineering, San Francisco, 2006, paper no. 2022. http://peer.berkeley.edu/nga/flatfile.html.

[31] Perlovsky LI. Neural networks and intellect. Oxford University Press; 2001.

[32] Cybenko J. Approximations by superpositions of a sigmoidal function. Math Control Signal Syst 1989;2:303–14.

[33] Alavi AH, Gandomi AH, Mollahasani A, Heshmati AAR, Rashed A. Modeling of maximum dry density and optimum moisture content of stabilized soil using artificial neural networks. J Plant Nutrit Soil Sci 2010;173(3):368–79.

[34] Haykins S. Neural networks – A comprehensive foundation. 2nd ed. Englewood Cliffs, NJ, USA: Prentice Hall International Inc.; 1999.

[35] Rumelhart DE, McClelland JL. Parallel Distributed Processing, vol. 1: Foundations. Cambrdige, MA: The MIT Press; 1986.

[36] Kirkpatrick S, Gelatt CD, Vecchi MP. Optimization by simulated annealing. Science 1983;220(4598):671–80.

[37] Cerny V. Thermodynamical approach to the traveling salesman problem: an efficient simulation algorithm. J Optim Theor Appl 1985;45:41–52.

[38] Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state calculations by fast computing mechanics. J Chem Phys 1953;21(6):1087–92.

[39] Aarts E. Simulated annealing and Boltzmann machines: a stochastic approach to combinatorial optimization and neural computing. NY, USA: Wiley; 1989.

[40] Alavi AH, Ameri M, Gandomi AH, Mirzahosseini MR. Formulation of flow number of asphalt mixes using a hybrid computational method. Construct Build Mater 2011;25(3):1338–55.

[41] Ingber L. Simulated annealing: practice versus theory. Math Comput Model 1993;18(11):29–57.

[42] Russel SJ, Norvig P. Artificial intelligence: a modern approach. 2nd ed. Englewood Cliffs: Prentice Hall; 2002.

[43] Reed RD, Marks RJ. Neural smithing: supervised learning in feedforward artificial neural networks. Cambridge: The MIT Press; 1999.

[44] Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Numerical recipes in C++: the art of scientific computing. 2nd ed. Cambridge: Cambridge University Press; 2002.

[45] Huang M, Romeo F, Sangiovanni-Vincentelli A. An efficient general cooling schedule for simulated annealing. In: Proceedings of the IEEE International Conference on Computer Aided Design, 1986, p. 381–384.

[46] Masters T. Practical neural network recipes in C++. London: Academic Press; 1993. p. 118–134.

[47] Jones MT. AI application programming. 2nd ed. Charles River Media; 2005.

[48] Luke BT. Simulated Annealing Cooling Schedules. 2007. Available at: http://members.aol.com/btluke/simanf1.htm.

[49] Mobius A, Neklioudov A, Diaz-Sanchez A, Hoffmann KH, Fachat A, Schreiber M. Optimization by thermal cycling. Phys Rev Lett 1997;79:4297–301.

[50] Rajabi AM, Khamehchiyan M, Mahdavifar MR, Del Gaudio V. Attenuation relation of Arias intensity for Zagros Mountains region (Iran). Soil Dyn Earthq Eng 2010;30:110–8.

[51] Mesbahi E. Application of artificial neural networks in modelling and control of diesel engines, PhD Thesis, University of Newcastle, UK, 2000.

[52] Goh ATC, Kulhawy FH, Chua CG. Bayesian neural network analysis of undrained side resistance of drilled shafts. J Geotech Geoenviron Eng 2005;131(1):84–93.

[53] Ledesma S. Neural-Lab, Version 3.1. University of Guanajuato, Mexico, 2009. Available at: http://www.dicis.ugto.mx/profesores/sledesma/documentos/index.htm.

[54] Maravall A, Gomez V. Eviews Software, Version 5, Quantitative Micro Software. Irvine C.A.: LLC; 2004.

[55] Smith GN. Probability and statistics in civil engineering. London: Collins; 1986.
[56] Frank IE, Todeschini R. The data analysis handbook. Amsterdam: Elsevier; 1994.
[57] Golbraikh A, Tropsha A. Beware of q2. J Mol Graph Modell 2002;20(4): 269–76.
[58] Kraslawski A, Pedrycz W, Nyström L. Fuzzy neural network as instance generator for case-based reasoning system: an example of selection of heat exchange equipment in mixing. Neural Comput Appl 1999;8(2):106–13.
[59] Garson GD. Interpreting neural-network connection weights. AI Expert 1991;6:47–51.