doi:10.1007/s11589-012-0843-5

Observations on the application of artificial neural network to predicting ground motion measures*

Hanping Hong♣ Taojun Liu and Chien-Shen Lee

Department of Civil and Environmental Engineering, University of Western Ontario London, Ontario N6A 5B9, *Canada*

Abstract Application of the artificial neural network (ANN) to predict pseudospectral acceleration or peak ground acceleration is explored in the study. The training of ANN model is carried out using feed-forward back-propagation method and about 600 records from 39 California earthquakes. The statistics of the residuals or modeling error for the trained ANN-based models are almost the same as those for the parametric ground motion prediction equations, derived through regression analysis; the residual or modeling error can be modeled as a normal variate. The similarity and differences between the predictions by these two approaches are shown. The trained ANN-based models, however, are not robust because the models with almost identical mean square errors do not always lead to the same predictions. This undesirable behaviour for predicting the ground motion measures has not been shown or discussed in the literature; the presented results, at least, serve to raise questions and caution on this problem. A practical approach to ameliorate this problem, perhaps, is to consider several trained ANN models, and to take the average of the predicted values from the trained ANN models as the predicted ground motion measure.

Key words: **neural network; peak ground acceleration; pseudospectral acceleration; seismic ground motion measures; uncertainty**

CLC number: P315.9 **Document code**: A

1 Introduction

Frequently used seismic ground motion measures for seismic hazard assessment are the ground acceleration (PGA), peak ground velocity (PGV), peak ground displacement (PGD) and pseudospectral acceleration (PSA). The ground motion predicting equations (GMPEs) (or attenuation relations) for these measures can be developed using historical records, parametric models and regression algorithms (Abrahamson and Youngs, 1992; Joyner and Boore, 1993; Hong et al., 2009). Many empirical equations have been developed (Singh et al., 1987; Ambraseys et al., 1996; Boore et al., 1997; Youngs et al., 1997; Campbell, 1997; Atkinson and Boore, 2003; Douglas, 2003; Hong and Goda, 2007). A special issue of *Earthquake Spectra* (February, 2008) presents GMPEs resulted from the Next Generation of Attenuation (NGA) project. These predicting equations have been used to assess seismic hazard and risk (Frankel et al., 1996; Adams and Halchuk, 2003; McGuire, 2004; Goda and Hong, 2009).

An alternative approach to predict the ground motion measures that use the artificial neural network (ANN) has also been considered. The ANN has three interconnected components: input layer, hidden layer(s) and output layer; the observed historical or experimental input and output information are used to train (synaptic) weights and biases to predict outputs (Rumelhart et al., 1986; Funahashi, 1989; Haykin, 1999; Principe et al., 1999). The application of the ANN to predict the ground motion measures has been presented by Wang (1993), Emami et al. (1996), Kerh and Chu (2002), Garcia et al. (2007), Ahmad et al. (2008), and Günaydin and Günaydin (2008). These s-

* **Received** 2 February 2012; **accepted in revised form** 29 March 2012; **published** 10 April 2012.

♣ **Corresponding author. e-mail**: Hongh@eng.uwo.ca

tudies considered ground motion parameters such as the PGA, PGV, and PGD. The studies showed that the ANN-based model could be a valuable alternative to process data and to predict the peak ground motion parameters because its predicted values are in good agreement with those obtained by using ground motion predicting equations, derived based on the same data set or taken directly from literature. In particular, Günaydin and Günaydin (2008) compared the application of three different ANN methods: the feed-forward back-propagation (FFBP), radial basis function, and generalized regression neural networks, and concluded that the FFBP performs better than the other two methods. However, none of the afore-mentioned studies considered the application of the ANN to predict the PSA. Other applications of the ANN for seismic ground motion predictions include the generation of artificial earthquake and response spectra (Lee and Han, 2002) and synthetic ground motion records (Lin and Ghaboussi, 2001; Amiri et al., 2009).

This study explores the application of the ANN model to predict the PSA. For the training of ANN and analysis, a set of 592 California records (from 39 earthquakes) extracted from the NGA database (PEER Center 2006) is employed. The input layer for the ANN considers the earthquake moment magnitude, closest horizontal distance to projected faults on the earth, focal depth, and site condition that is defined by the average shear wave velocity in the uppermost 30 m, as these parameters are often considered for developing the GMPEs. The feed-forward back-propagation (FFBP) is employed to search the weights and biases for the ANN model. Comparison of the predicted PSA obtained by using the ANN and the GMPEs are presented. For the study to be self-contained, the basic concepts of the ANN are summarized in the following section. This is followed by the description of the considered ground motion records, the analysis results obtained by using the FFBP, comparison of the predicted ground motion measures using the ANN based model and those obtained from the GMPEs and conclusions.

2 Basic concepts of artificial neural network

An ANN is a mathematical model or computational model; it attempts to mimic the structure and/or functional aspects of biological neural networks (Rumelhart et al., 1986; Funahashi, 1989; Haykin, 1999). The ANN has three interconnected components: input layer, hidden layer(s) and output layer, each layer with one or more neurons (nodes). It is a non-linear data modeling tool that can be used to model complex relationships between inputs and outputs; one does not have to understand the causality to apply the ANN model.

A typical feed-forward back-propagation network is illustrated in Figure 1. Each neuron in the network receives information from its predecessors, weighs them and transfers them into output value(s). In particular, if there is only one hidden layer, the output of the k-th neuron in the output layer, denoted by y_k, can be expressed in the following mathematical form,

$$y_k = h\left\{\sum_{j=1}^{m} \boldsymbol{w}_{kj,2} f\left[\sum_{i=1}^{n} (\boldsymbol{w}_{ji,1} x_i) + \boldsymbol{\theta}_{j,1}\right] + \boldsymbol{\theta}_{k,2}\right\}, \quad k = 1,\ 2,\ \cdots,\ n_0, \tag{1}$$

where n_0 is the total number of output neurons, n is the total number of neurons in the input layer, $\boldsymbol{w}_{ji,1}$ are the synaptic weights between the input and hidden layers, $\boldsymbol{w}_{kj,2}$ are the synaptic weights between the hidden and output layers, $\boldsymbol{\theta}_{j,1}$ and $\boldsymbol{\theta}_{k,2}$ are the biases associated with the hidden and output layers, $f(\ )$ is the transfer (or activation) function between the input and hidden layers, $h(\ )$ is the transfer (or activation) function between the hidden and output layers, and m is the total number of neurons in the hidden layers.

Commonly used transfer functions include the tan-sigmoid, log-sigmoid and linear functions. For example, Emami et al. (1996) and Garcia et al. (2007) adopted the sigmoid function, Ahmad et al. (2008) used the

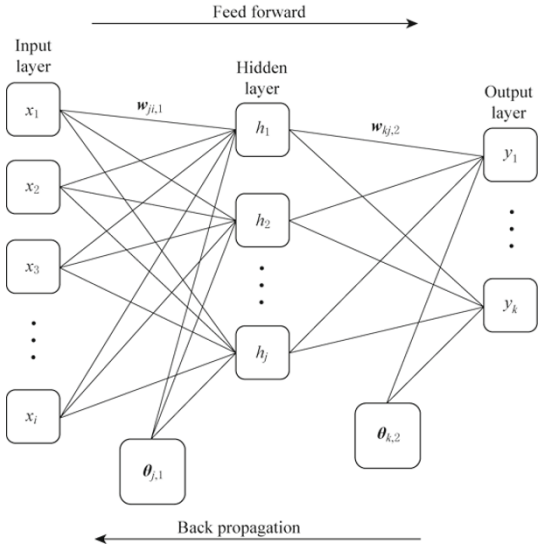

Figure 1 Schematic diagram of artificial neural network model (feed-forward back-propagation).

tan-sigmoid function, Günaydin and Günaydin (2008) considered combinations of tan-sigmoid and log-sigmoid functions for the estimation of the peak ground motion parameters.

Variety of learning techniques can be used to evaluate the weights and biases including the back-propagation (Rumelhart et al., 1986; Haykin, 1999), where the minimization of a predefined error function, such as the sum of the square of the differences between the predicted and observed outputs, (i.e., mean square error (MSE)), is carried out. In this study, training of weights and bias for the ANN model, by minimizing the MSE is carried out using the Levenberg-Marquardt algorithm (Marquardt, 1963; Press et al., 1992), implemented in Matlab (Hagan and Menhaj, 1994).

3 Strong ground motion records

To develop the ANN model for predicting the ground motion measures, a set of 592 California records (from 39 earthquakes) extracted from the NGA database (PEER Center 2006) is adopted. The selection criteria and the detailed selected records are given in Hong and Goda (2007), and their GMPEs are summarized in Table 1 for a few selected vibration periods. The records are from earthquakes with moment magnitude, M, greater than 5.0, and for the NEHRP site class B, C, or D with the Geomatrix's classification I, A, or B (see PEER Center for details). These records are band-pass filtered with the high-pass corner frequency less than 0.5 Hz.

Table 1 Ground motion predicting equations and their coefficients for selected vibration periods (Hong and Goda, 2007)

Ground motion measure	T_n/s	b_1	b_2	b_3	b_4	b_5	h/km	σ_T
PGA	–	1.096	0.444	0.0	−1.047	0.038	5.7	0.501
PGA_{ref}	–	0.851	0.480	–	−0.884	–	6.3	–
PSA	0.2	2.171	0.305	−0.045	−1.158	0.061	8.0	0.534
	0.5	1.405	0.965	−0.001	−0.746	−0.073	5.7	0.608
	1.0	0.481	0.644	−0.229	−0.843	−0.002	4.9	0.694
	1.5	0.114	1.044	−0.298	−0.566	−0.124	4.1	0.692

Note: Ground motion predicting equation: $\ln Y = b_1 + b_2(M-7) + b_3(M-7)^2 + (b_4 + b_5[M-4.5])\ln[(D^2+h^2)^{0.5}] + AF_s$, where Y represents PSA or PGA, M is the moment magnitude of the earthquake, D(km) is the closest horizontal distance from the station to a point on the Earth's surface that lies directly above the rupture, h (km) represents a fictitious depth, and AFs represents the amplification factor due to linear and nonlinear soil behavior suggested by Atkinson and Boore (2006), which is a function of the shear-wave velocity v_S and the (expected) PGA at the reference soil condition PGA_{ref} for v_S equal to 760 m/s. The adopted relation to calculate the PGA_{ref} is given as, $\ln PGA_{ref} = b_1 + b_2(M-7) + b_4 \ln[(D^2+h^2)^{0.5}]$. The PSA and PGA are in fraction of gravitational acceleration (g).

4 ANN model for predicting ground motion measures

4.1 Impact of number of neurons in the hidden layer

Throughout this study, for simplicity and to avoid potential over fitting because of limited number of records, only a single hidden layer is considered; the four neurons in the input layer are used to represent the earthquake moment magnitude M, closest horizontal distance to projected faults on the Earth D (km), focal depth h (km), and the average shear wave velocity in the uppermost 30 m, v_{S30} (m/s). The output represents the logarithmic of ground motion measure (i.e., lnPSA), because the GMPEs are expressed as logarithmic of PSA or PGA, and its use could reduce nonlinearly between inputs and output. For parametric assessment, first, we randomly select 80% of the 592 records and their associated PSA at the vibration period, T_n, equal to 0.2 (s) and the damping coefficient ξ equal to 5%. For each record, the two PSA values corresponding to the two horizontal record components are used for the training purpose.

By considering three neurons in the hidden layer, and adopting the tan-sigmoid function,

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{2}$$

as the transfer function between the input and hidden layers, and the linear function as the transfer hidden and output layers, the solution of the weights and biases are obtained by using the algorithm implemented in Matlab (Hagan and Menhaj, 1994) with a set of randomly generated initial weights and biases using the Nguyen-Widrow method (Nguyen and Widrow, 1990).

For the analysis, the adopted convergence criteria are that the performance gradient is less than 1.0×10^{-5}, or MSE equals to zero (use of performance gradient less than 1.0×10^{-10} is also considered and the obtained results are similar to those by using 1.0×10^{-5}, they are not mentioned further).

To investigate whether the obtained solution corresponds to a local or global minimum MSE, the aforementioned analysis is repeated 200 times (i.e., 200 trial, each trial with sufficient number of learning epochs or of step in the learning phase until the adopted convergence criteria is achieved), but for each trial a new set of randomly generated initial weights and biases are employed, the obtained MSE for each trial (i.e., each obtained solution) is shown in Figure 2. It was observed that for each trial the required number of learning epochs to achieve convergence is less than 3000. The results indicate that the MSE surface for the present problem is complex and may contain many local minima; the MSE for more than 85% cases differs from the lowest MSE (that corresponds to Trial No. 188) by less than 10%. By repeating this process five times, each with 200 trials, it was observed that in all cases the difference between the lowest MSE obtained from 200 trials and that of Trial No. 188 is less than 1%, indicating that the use of 200 trials to identify the solution may be adequate. However, if the trained models (each using 200 trials) are used to predict the PSA, it was observed that the trained ANN models in this manner do not always provide consistent predictions. Although, an attempt is made by increasing the number of trials

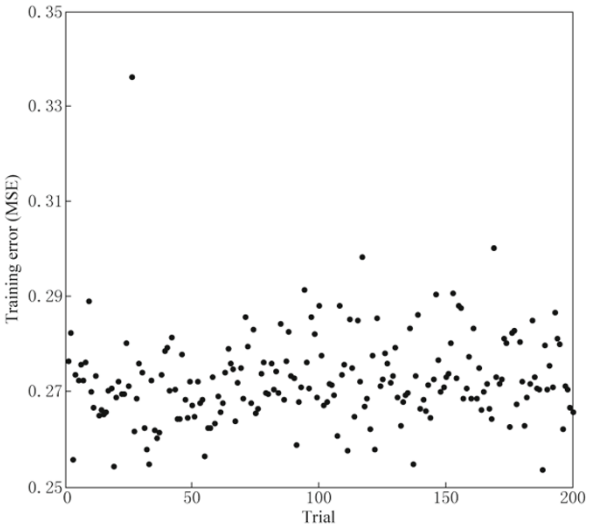

Figure 2 Mean square error (MSE) for each trial considering PSA at T_n=0.2 s from 80% of records.

up to 1 000 to find the solution, it does not eliminate the observed problem. In other words, the results suggest that the trained ANN model for the considered records should be not robust. The implication of this will be discussed later again.

If the results from Trial No. 188 are used as the solution, samples of the modeling error versus M, D, h and v_{S30} are plotted in Figures 3a to 3d, where the modeling error is defined as the difference between the predicted output and the corresponding observed (or target) output [in this case it represents ln(PSA) at 0.2 (s)]. The plot indicates that the modeling error is not correlated with the input parameters. Furthermore, visual inspection for the samples of the modeling error plotted on the normal probability in Figure 3e indicates that the error could be modeled as a normal variate with mean of practically zero and standard deviation of 0.503. Also, the Kolmogorov-Smirnov goodness-of-fit test (Benjamin and Cornell, 1970) indicates that the normality hypothesis could not be rejected at a significance level of 10%. The probabilistic characterization of this modeling error is important because this uncertainty must be incorporated in the ANN-based model to assess the seismic hazard and risk. This is similar to the case when the GMPEs are used for such purpose (Goda and Hong 2009). The standard deviation shown in Figure 3e represents the total variation; it accounts for the interevent variability, intraevent variability and the random orientation variability (Boore et al., 1997). The predicted PSA by using the trained ANN model for the remaining 20% of records is illustrated in Figure 3f and compared with the observed PSA; the modeling error associated with the prediction shown in Figure 3f is plotted in Figure 3g. The results suggest that the trained ANN model may be a useful tool to predict the PSA.

A comparison of the mean and standard deviation of the modeling error by considering each trial depicted in Figure 2 as the solution is presented in Figures 4a and 4b. The figures indicate that the statistics of the modeling error for a trial with a MSE, which differs only by up to 5% from the lowest MSE, do not differ significantly from that shown in Figure 3e. This supports the selection of solution strategy mentioned earlier (i.e., carried out several trials, and select the results with the lowest MSE as the solution). Therefore, for the remaining analysis in this study, the solution for each case is obtained based on 200 trials.

The ANN model depends on the adopted transfer functions. Use of log-sigmoid function in combination with tan-sigmoid or linear functions is also considered.

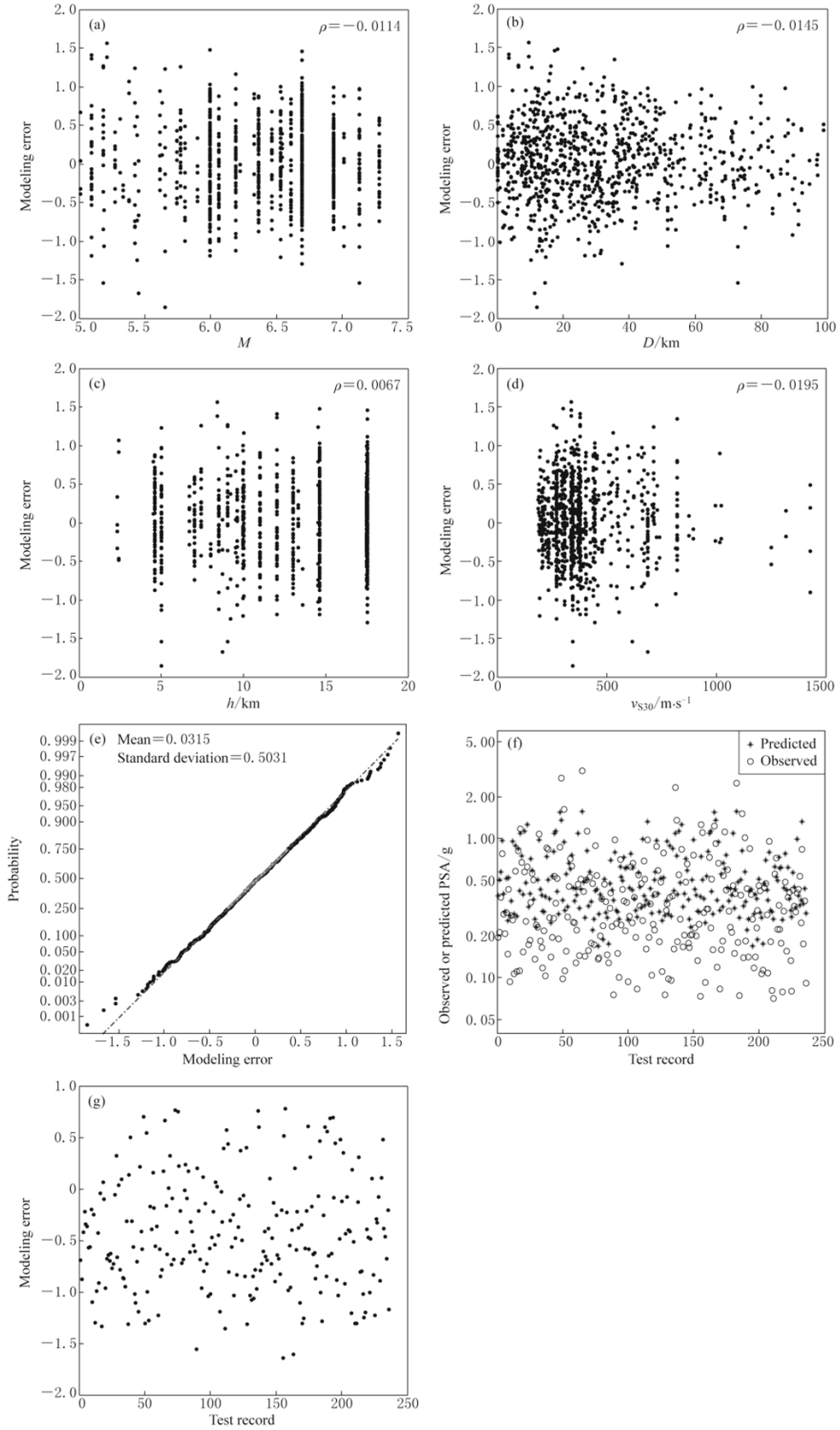

Figure 3 Modeling error associated with the trained model (ρ in the plots denotes the correlation coefficient): (a) to (d) Modeling error versus M, D, h, and v_{S30}, (e) Modeling error depicted on normal probability paper, (f) Comparison of predicted versus observed values for test records, (g) Modeling error of test records (i.e. 20% of records that are not used for training ANN model).

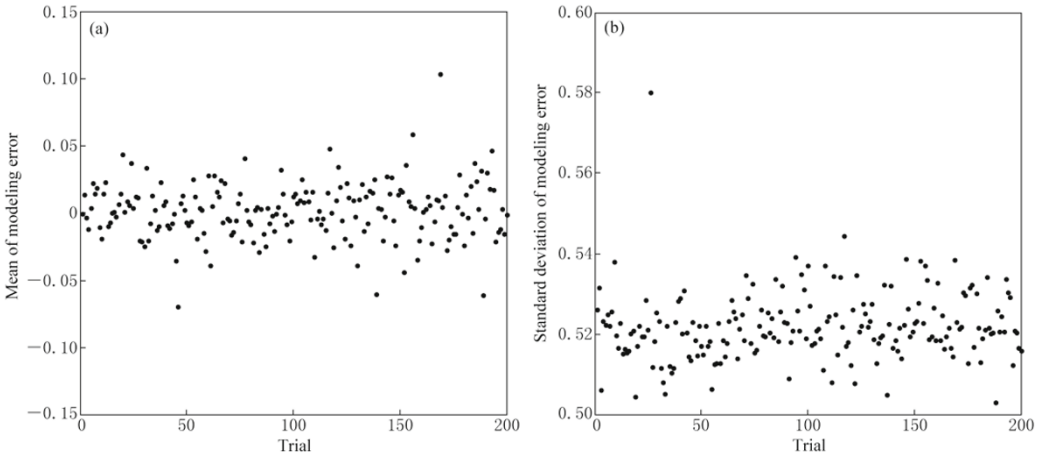

Figure 4 Comparison of statistics of the modeling error (for randomly selected 80% of all records) associated with different trials: (a) Mean, (b) Standard deviation.

As the results of these combinations and those showed in Figure 3 for the considered case indicate that the use of both tan-sigmoid and linear functions leads to the lowest MSE, this combination of transfer functions is used for the remaining numerical analysis. It must be emphasized that this observation (i.e., the most adequate combination of the transfer functions) may not be generalized to other sets of records.

Now, as indicated previously, the adequacy of the ANN model depends on the number of neurons in the hidden layers. An increase in the number of neurons might result in a reduced MSE but drastically increases the number of weights and biases that need to be trained. To find the impact of the number of neurons in the hidden layer on the MSE of the trained ANN model, we repeat the analysis carried out previously, by considering the number of neurons equal to 2, 5, 10 and 15. The MSE and the empirical probability distribution of the modeling error associated with the selected solution (based on 200 trials as discuss earlier) are shown in Figure 5. Results shown in Figures 2, 3e and 5 indicate that the MSE tends to decrease as the number of neurons increases, and the standard deviation of the modeling error decreases slightly as the number of neurons increase. Based on the results shown in Figure 5 and the consideration that a model with fewer parameters is often preferred in practical application, the use of only three neurons in the hidden layer is considered for the remaining part of this study.

4.2 Effect of percent of records used for training

To illustrate the effect of percent of records used for training on the trained ANN model, we considered this percentage value equal to 75%, 85% and 95%. The modeling error associated with each trained ANN model is presented on the normal probability paper shown in Figure 6. As the records are randomly selected from the considered 592 records, the standard deviation of the modeling error shown in the figure only reflects a particular set of selected records. The plotted results suggest that in all cases the modeling error can be adequately modeled as normal variate. The standard deviation varies from 0.508 to 0.510, which is lower than 0.534 that is obtained based on GMPEs and regression analysis as shown in Table 1. The fact that the statistics of the modeling error for the ANN model and for the GMPE are similar indicates that the development of ANN-based model for predicting the PSA should be explored. This development is described below by using all the considered records.

4.3 Trained ANN models for predicting PSA and PGA

To develop a set of ANN models to predict the ground motion measures, the PGA and PSA at T_n equal to 0.2, 0.5, 1.0 and 1.5 (s) for the 592 records are employed. For each ground motion measure, the obtained MSE values are shown in Figure 7. The trained weights and biases of the models are listed in Table 2, and the minimum and maximum values of the input and target output parameters used in the scaling functions to define the normalized input and output parameters for training are listed in Table 3. Samples of the modeling error versus M, D, h and v_{S30} are depicted in Figure 8, suggesting that the error is not significantly correlated with M, D, h and v_{S30}. Furthermore, the plots of the samples of the modeling error shown on the normal probability paper in Figure 9 indicate that the modeling

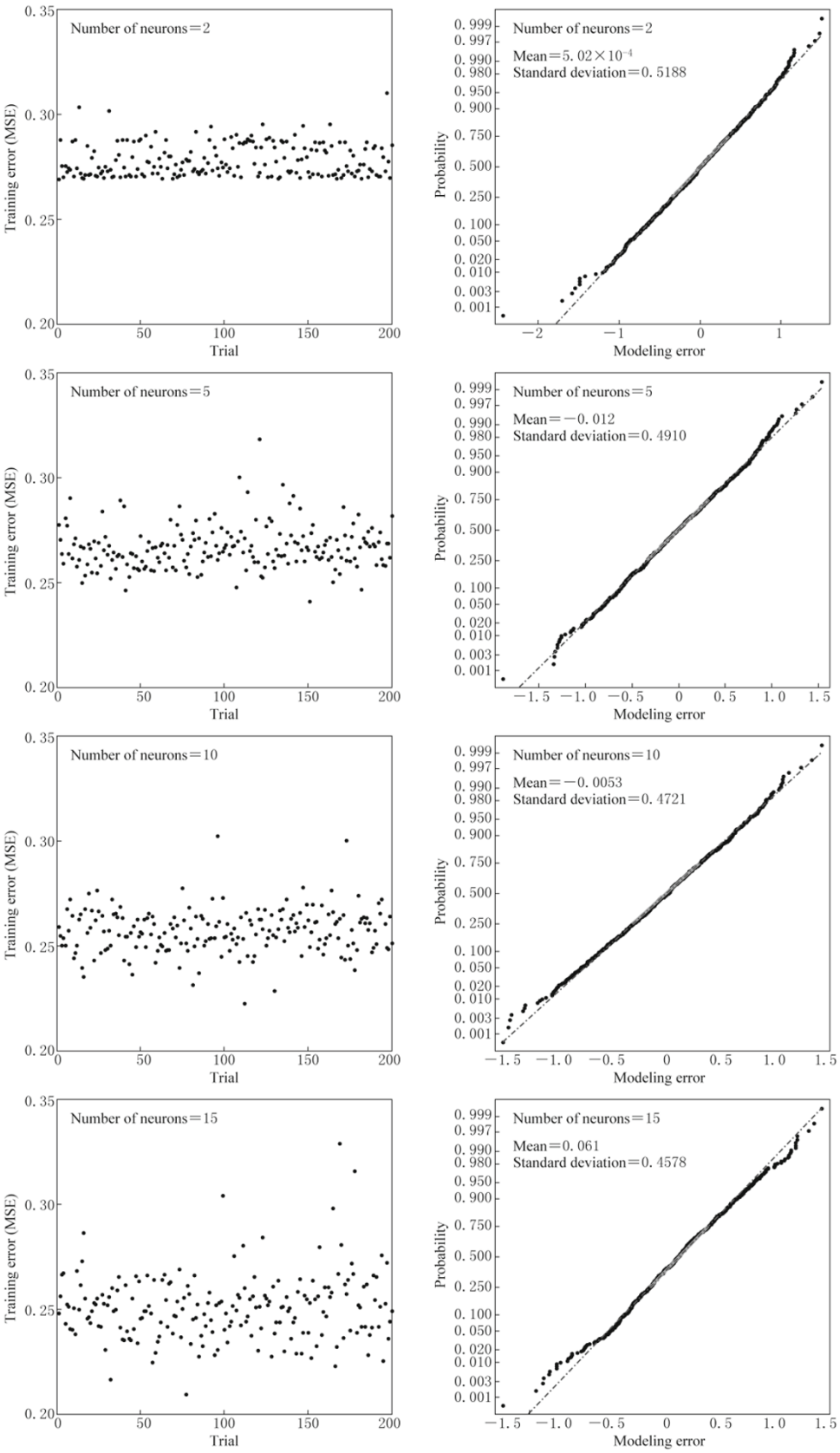

Figure 5 Mean square error for different trials and modeling error corresponding to the lowest MSE considering 2, 5, 10 and 15 neurons in the hidden layer.

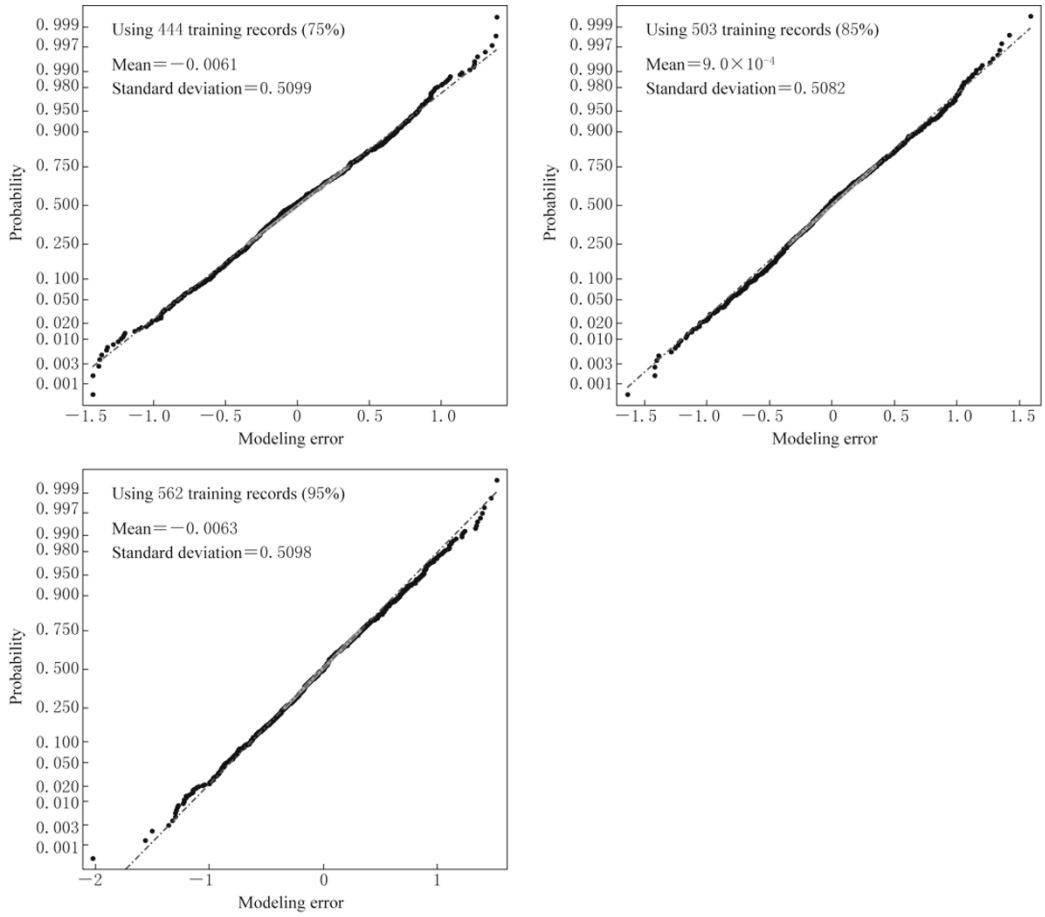

Figure 6 Samples of modeling error for models trained based on 75%, 85%, and 95% of all the considered records.

Table 2 Trained ANN based prediction models (the weights and bias are to be applied with the normalized inputs defined in Table 3)

Ground motion measure	Input weights $\boldsymbol{w}_{ji,1}$	Hidden layer weights $\boldsymbol{w}_{kj,2}$	Biases $\boldsymbol{\theta}_{j,1}$	Biases $\boldsymbol{\theta}_{k,2}$
PGA	$\begin{bmatrix} -0.03943 & -2.73202 & 0.058007 & 0.340969 \\ -12.0485 & -0.88578 & -8.84292 & 0.447709 \\ -0.39879 & 0.267677 & -0.1302 & 0.282241 \end{bmatrix}$	$\begin{bmatrix} 0.362403 & 0.10599 & -0.99579 \end{bmatrix}$	$\begin{bmatrix} -1.81987 \\ -3.01071 \\ -0.00958 \end{bmatrix}$	$[-0.24545]$
PSA at 0.2 s	$\begin{bmatrix} 2.44444 & 8.647368 & 6.773296 & -0.78949 \\ 0.165255 & -2.62223 & 0.381331 & 0.185042 \\ -0.1556 & 0.14367 & -0.05507 & 0.131044 \end{bmatrix}$	$\begin{bmatrix} -0.11884 & 0.299766 & -2.62301 \end{bmatrix}$	$\begin{bmatrix} 7.886026 \\ -1.98237 \\ -0.55432 \end{bmatrix}$	$[-1.42504]$
PSA at 0.5 s	$\begin{bmatrix} -1.13569 & -4.42034 & 1.754484 & 0.148884 \\ -0.2254 & -0.0964 & -0.52287 & 0.012732 \\ -0.17069 & -0.01262 & -0.28914 & 0.038621 \end{bmatrix}$	$\begin{bmatrix} 0.178388 & 7.317167 & -18.0766 \end{bmatrix}$	$\begin{bmatrix} -3.04989 \\ -0.56503 \\ -0.83166 \end{bmatrix}$	$[-8.94124]$
PSA at 1.0 s	$\begin{bmatrix} 0.328156 & -0.32842 & -0.32408 & -0.63991 \\ 0.130522 & -0.59592 & 0.802454 & 0.431288 \\ -3.24718 & -1.12761 & -3.43098 & 0.182526 \end{bmatrix}$	$\begin{bmatrix} 1.095043 & 0.7011 & -2.1957 \end{bmatrix}$	$\begin{bmatrix} -0.91157 \\ -1.04963 \\ 7.361045 \end{bmatrix}$	$[2.979618]$
PSA at 1.5 s	$\begin{bmatrix} -1.00287 & 6.382422 & 0.17595 & -2.3223 \\ 0.893689 & 0.275995 & 2.989585 & 0.227485 \\ 0.924012 & -0.93933 & -0.40178 & -1.28787 \end{bmatrix}$	$\begin{bmatrix} -0.18216 & 5.350651 & 0.546571 \end{bmatrix}$	$\begin{bmatrix} 4.17248 \\ -5.17299 \\ -1.27318 \end{bmatrix}$	$[5.391307]$

error can be modeled as normal variate. In all cases, the Kolmogorov-Smirnov goodness-of-fit test (Benjamin and Cornell, 1970) indicates that the normality hypothesis could not be rejected at a significance level of 10%. It must be emphasized that the standard deviation shown in Figure 9 represents the estimated standard deviation of the natural logarithmic of PSA or PGA, and that it includes the interevent, intraevent

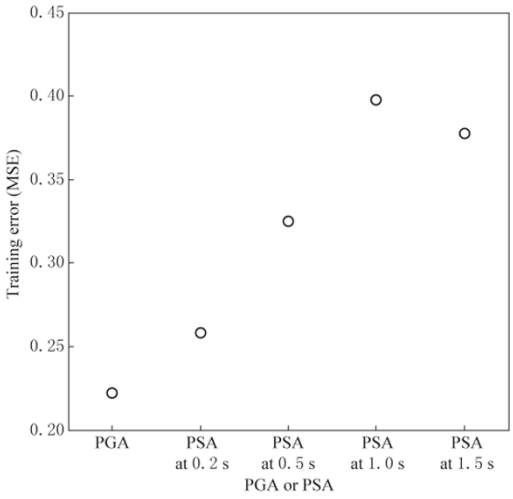

Figure 7 Mean square error for the trained ANN model to predict PGA or PSA.

Table 3 Minimum and maximum values of the input and target output parameters used in defining the normalized input and output parameters for training

	Minimum	Maximum
M	5.01	7.28
D/km	0	98.83
H/km	2.3	17.5
v_{S30}/m·s^{-1}	184.75	1428
ln(PSA) at 0.2 s	−3.364	1.118
ln(PSA) at 0.5 s	−4.482	1.094
ln(PSA) at 1.0 s	−5.039	0.602
ln(PSA) at 1.5 s	−5.781	−0.011
ln(PGA)	−3.925	0.576

Note: The input or target output x is normalized using $x_n=2\times(x-x_{\min})/(x_{\max}-x_{\min})-1$, where x_n is the normalized quantity of x, and $x_{\min}$ and $x_{\max}$ denote the minimum and maximum values of x observed from the training dataset. The PSA and PGA are in fraction of gravitational acceleration (g).

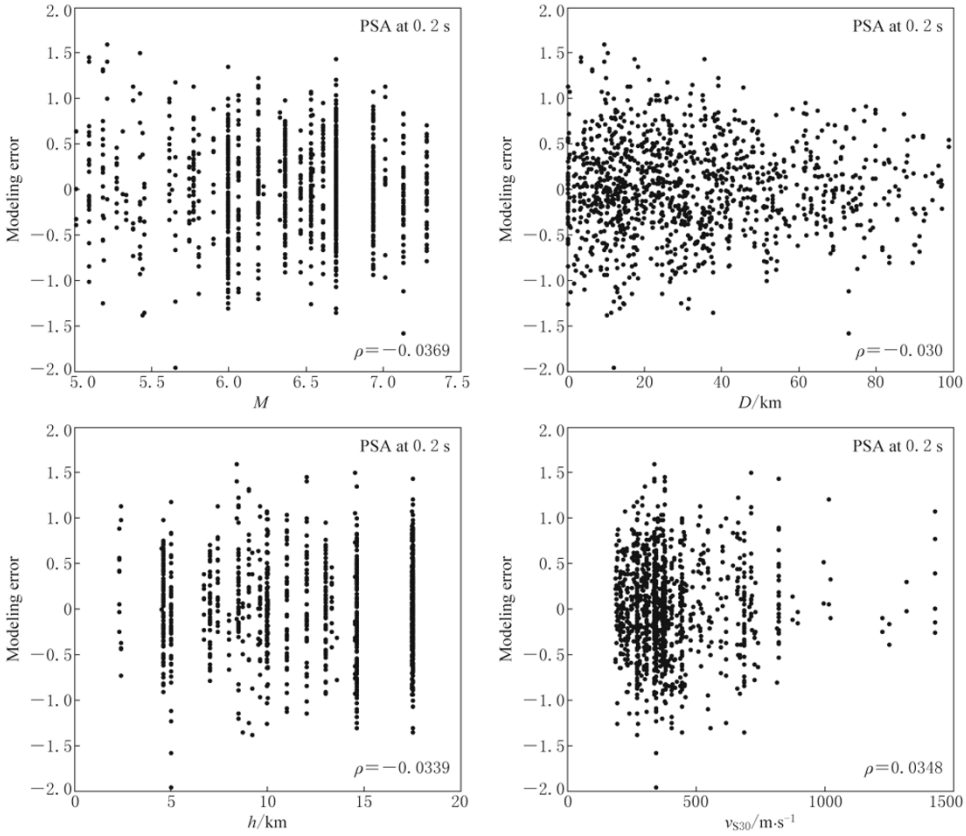

Figure 8 Plots of modeling error versus M, D, h and v_{S30}, for PSA at different vibration periods and PGA.

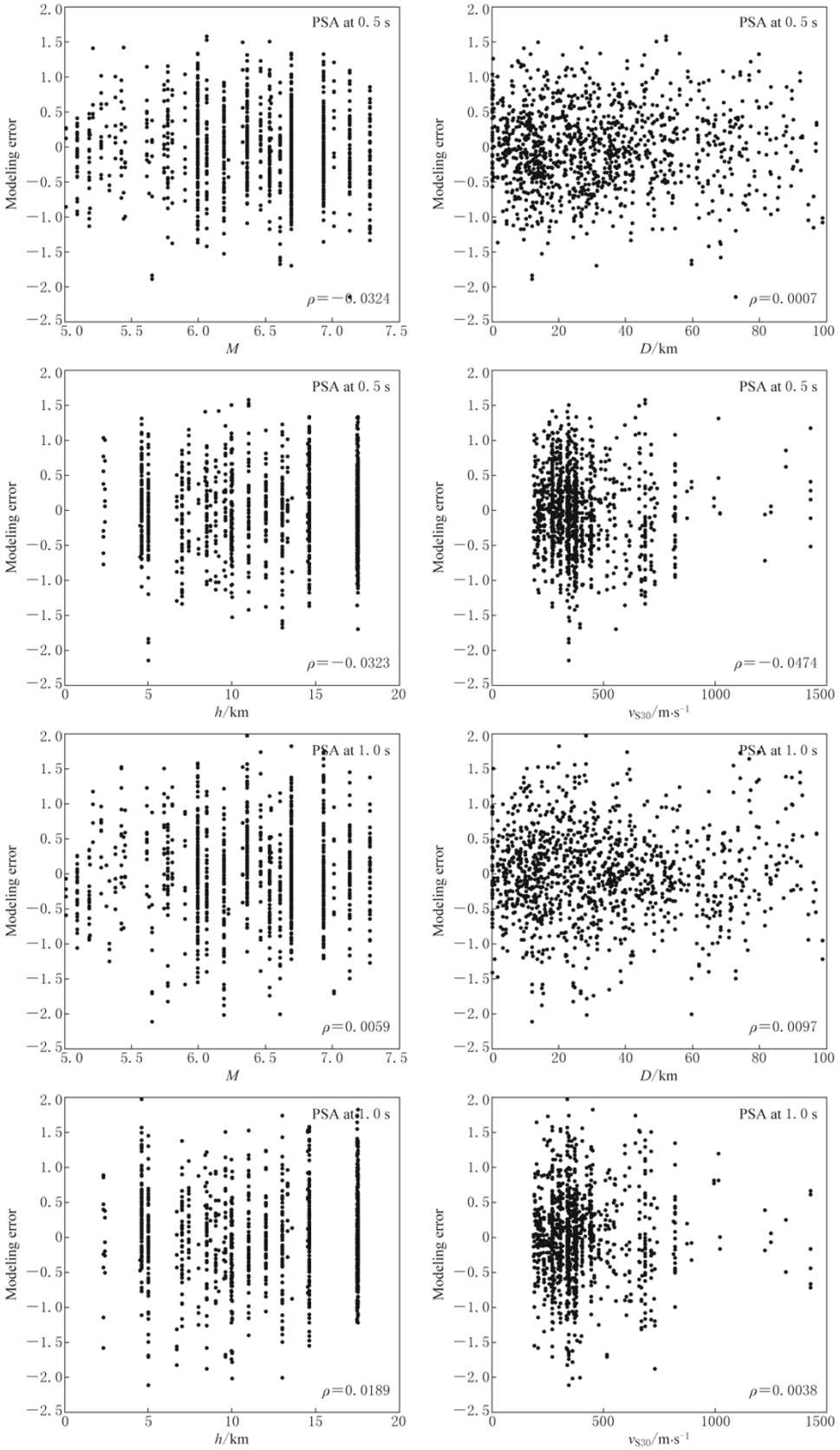

Figure 8 Plots of modeling error versus M, D, h and v_{S30}, for PSA at different vibration periods and PGA.

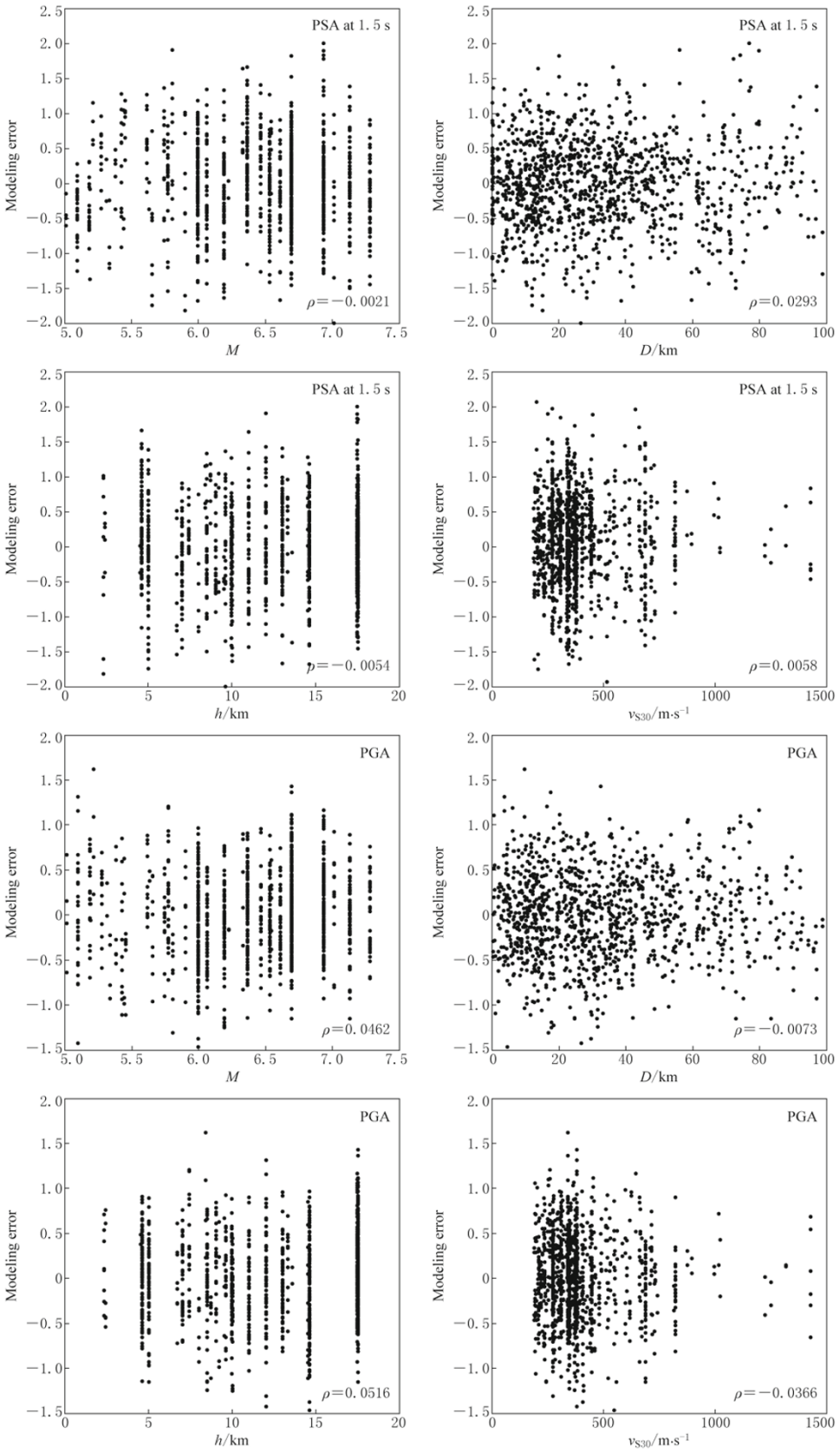

Figure 8 Plots of modeling error versus M, D, h and v_{S30}, for PSA at different vibration periods and PGA.

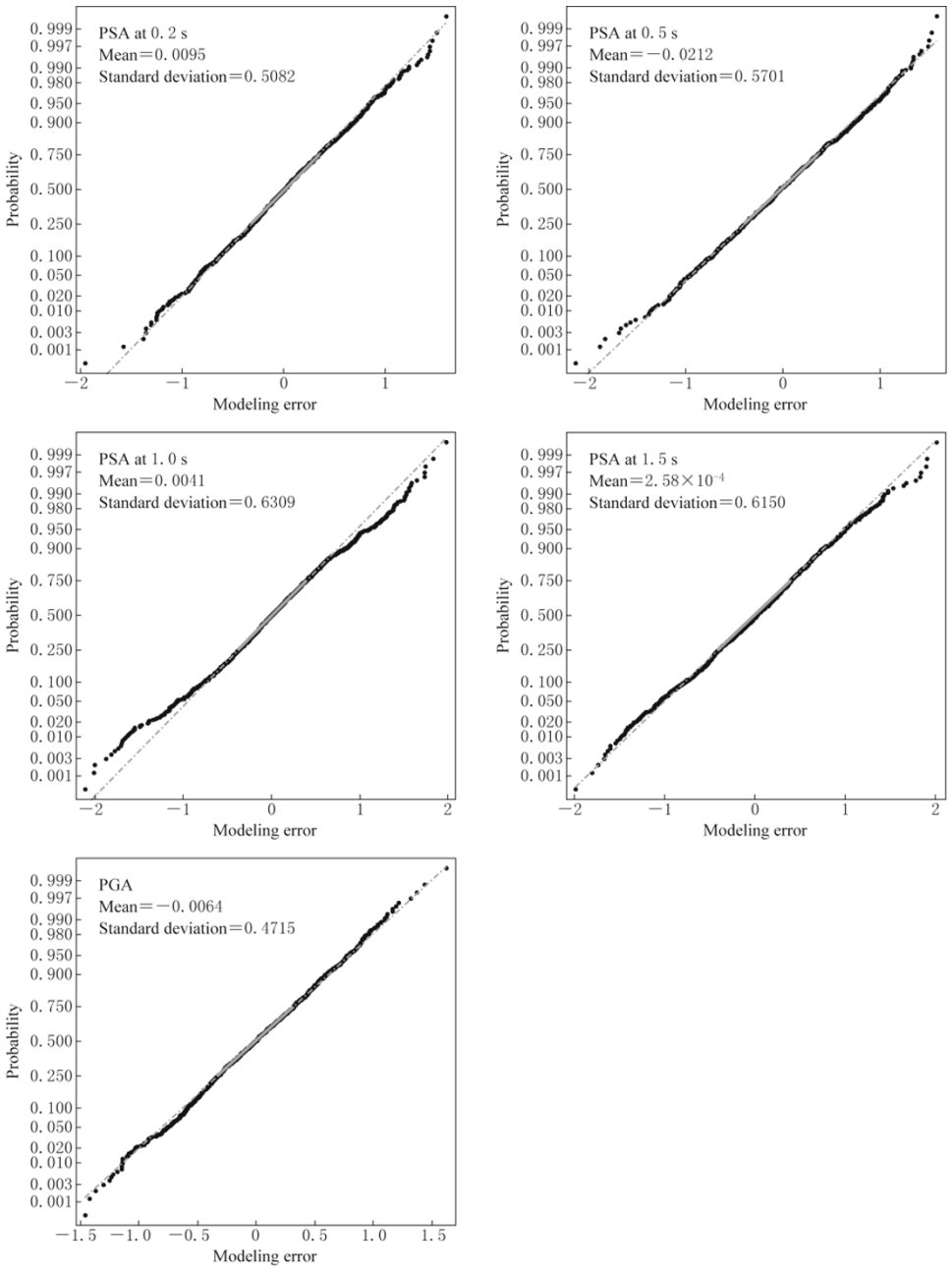

Figure 9 Normal probability plots of modeling error for PSA at different vibration periods and PGA.

and random orientation variability. These standard deviations are comparable to those shown in Table 1, which are obtained based on the regression analysis and using parametric ground motion prediction models, indicating that the accuracy of the ANN-based predicting models to the "observed" samples is comparable to the GMPEs listed in Table 1.

To further inspect the adequacy of the ANN-based predicting models, a comparison of the predicted ground motion measures by these models to those obtained by using GMPEs shown in Table 1 is depicted in Figure 10. It can be observed from the figure that the predicted ground motion measures by using trained ANN models are similar to those predicted by the GMPEs, although differences between the predicted values by the two approaches are visible. In general, the differences between the predicted ground motion measures decrease as D decreases; there is no indication whether the differences depend on the natural vibration period, or are functions of the PGA or PSA.

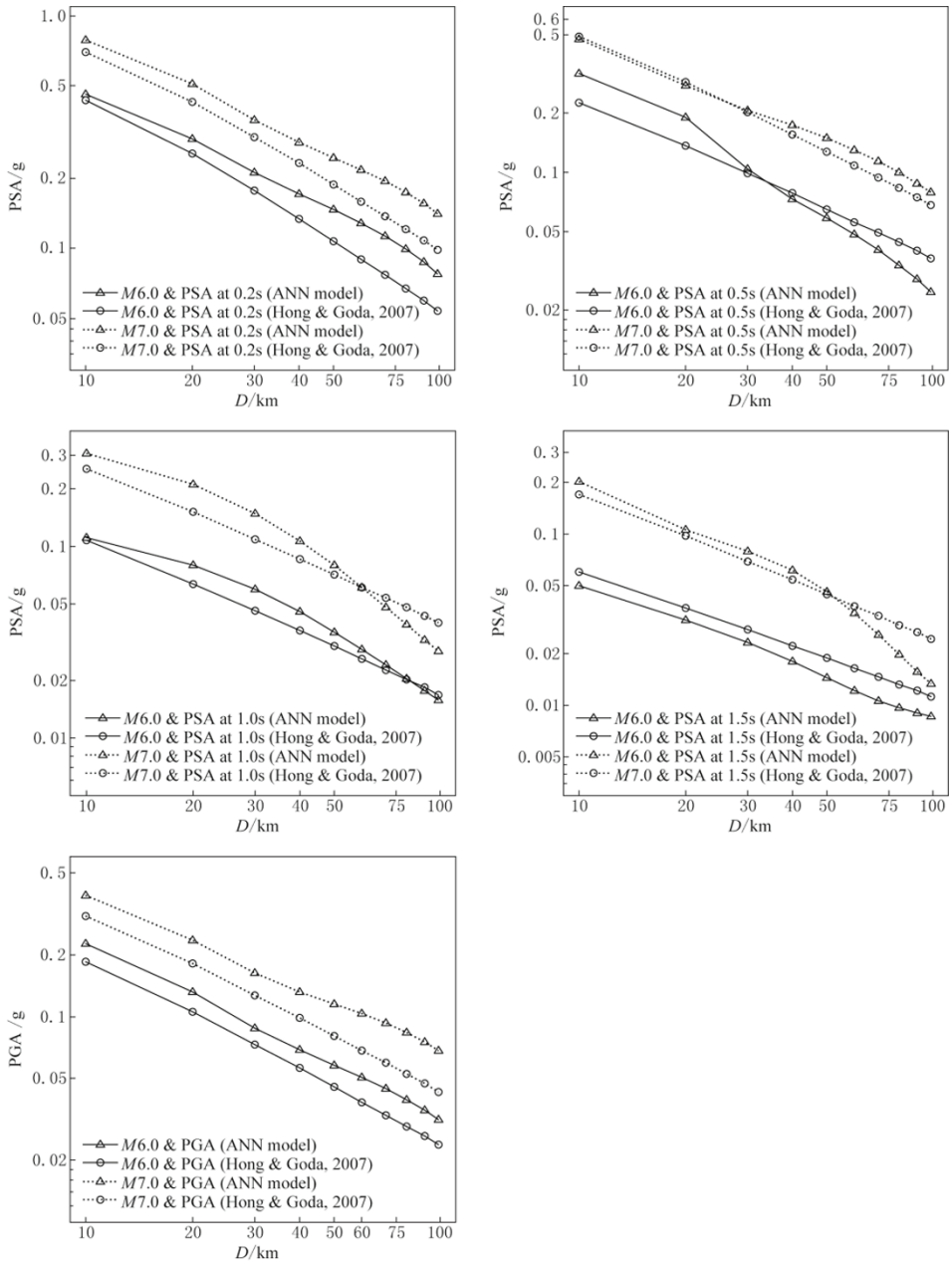

Figure 10 Comparison of predicted PSA and PGA by ANN models and by GMPEs for h=12.4 km and v_{S30}=394 m/s.

For the results presented in Figures 7 to 10, and Table 2, the trained ANN model is always selected from the trial with the lowest MSE out of 200 trials. As mentioned earlier, the trained ANN models, each with 200 trials, do not always provide very consistent predictions as shown in Figure 11, although in all cases the relative absolute differences between the MSE of any two trained models are less than 1%, and the standard deviation is almost identical (with relative absolute difference less than about 0.5%). A practical approach to overcome this problem, perhaps, is to consider several trained ANN-models, and to take the average of the predicted values from the trained ANN-models as the predicted ground motion measures.

5 Discussion and summary

Numerical experiments are conducted to investigate whether the ANN-based model could adequately predict the ground motion measure. For the analysis,

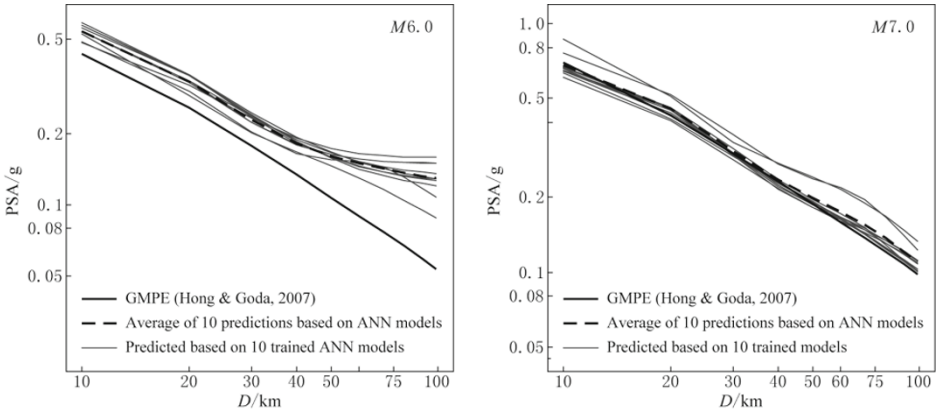

Figure 11 Comparison of predicted PSA at 0.2 (s) by several trained ANN models (each trained based on 200 trials) and by GMPEs for h=12.4 km and v_{S30}=394 m/s.

the feed-forward back-propagation method is considered and 1184 record components (i.e., 592 records) from California earthquakes are considered.

In all cases, the modeling error for the trained ANN-based models is very similar to that for the parametric ground motion prediction equations (GMPEs) obtained through regression analysis. It is shown that the modeling error for the ANN-based model can be modeled as a normal variate, and that there is negligible (linear) correlation between the modeling error and the input parameters: the earthquake moment magnitude, closest horizontal distance to projected faults on the earth, focal depth, and site condition. These indicate that the performance of the ANN-based model is comparable to that of the parametric ground motion prediction equations, if the minimum modeling error alone is used as the performance measure.

The trained ANN prediction models for the considered records, however, is not robust because the trained models with almost identical mean square errors do not always lead to the same predicted PSA or PGA. This undesirable behaviour for predicting the ground motion measures has not been shown or discussed in the literature; the presented results, at least, serve to raise questions and caution on this problem. If the use of the ANN-based model is preferred (as one does not have to understand the causality to apply the ANN model), a practical approach that may be used to ameliorate this problem is to take the average of the predicted values from several trained ANN models as the predicted ground motion measure. It must be emphasized that similar to the application of the GMPEs, the modeling error for ANN-based prediction model must be incorporated in assessing the seismic hazard and risk.

Acknowledgements The acknowledgements was included in a file entitled “Title Page” by your requirements for submission. If you could not find such a file please let us know. For your convenience we also included a file containing “acknowledgements” in this reply. Note that there is no number associated with “The financial support received from the Natural Science and Engineering Research Council of Canada and the University of Western Ontario is gratefully acknowledged” but there is a number associated with the scholarship for TJL.

References

Abrahamson N A and Yongs R R (1992). A stable algorithm for regression analyses using the random effects model. *Bull Seismol Soc Am* **82**(1): 505–510.

Adams J and Halchuk S (2003). Fourth generation seismic hazard maps of Canada: values for over 650 Canadian localities intended for the 2005 National Building Code of Canada. Open-File 4459, Geological Survey of Canada, Ottawa, Ontario, Canada.

Ahmad I, El Naggar M H and Khan A N (2008). Neural network based attenuation of strong motion peaks in Europe. *Journal of Earthquake Engineering* **12**(5): 663–680.

Ambraseys N N, Simpson K A and Bommer J J (1996). Prediction of horizontal response spectra in Europe. *Earthquake Engineering and Structural Dynamics* **25**(4): 371–400.

Amiri G G, Bagheri A and Razaghi S A S (2009). Generation of multiple earthquake accelerograms compatible with spectrum via the wavelet packet transform and stochastic neural networks. *Journal of Earthquake Engi-*

neering **13**(7): 899–915.

Atkinson G M and Boore D M (2003). Empirical ground-motion relations for subduction-zone earthquakes and their application to Cascadia and other regions. *Bull Seismol Soc Am* **93**: 1 703–1 729.

Benjamin J R and Cornell C A (1970). *Probability, Statistics and Decision for Civil Engineers*. McGraw-Hill, New York.

Boore D M, Joyner W B and Fumal T E (1997). Equations for estimating horizontal response spectra and peak acceleration from western North America. *Seism Res Lett* **68**: 128–153.

Campbell K W (1997). Empirical near-source attenuation relationships for horizontal and vertical components of peak ground acceleration, peak ground velocity, and pseudo-absolute acceleration response spectra. *Seism Res Lett* **68**(1): 154–179.

Douglas J (2003). Earthquake ground motion estimation using strong-motion records: A review of equations for the estimation of peak ground acceleration and response spectral ordinates. *Earth-Science Reviews* **61**(1–2): 43–104.

Emami S M R, Harada T and Iwao Y (1996). Prediction of peak horizontal acceleration using an artificial neural network model. *Structural Engineering and Earthquake Engineering* JSCE **13**(2): 111s–118s.

Frankel A, Mueller C, Barnhard T, Perkins D, Leyendecker E V, Dickman N, Hanson S and Hopper M (1996). National seismic hazard maps. Open-File 96-532, U.S. Department of the Interior, U.S. Geological Survey, Denver, CO, pp. 5–8.

Funahashi K (1989). On the approximate realization of continuous mappings by neural networks. *Neural Networks* **2**(3): 183–192.

Garcia S R, Romo M P and Mayoral J M (2007). Estimation of peak ground accelerations for Mexican subduction zone earthquakes using neural networks. *Geofisica Internacional* **46**(1): 51–63.

Goda K and Hong H P (2009). Deaggregation of seismic loss of spatially distributed buildings. *Bull Earthq Eng* **7**: 255–272.

Günaydin K and Günaydin A (2008). Peak ground acceleration prediction by artificial neural networks for northwestern Turkey. *Mathematical Problems in Engineering* Art. No.: 919420

Hagan M T and Menhaj M (1994). Training feed-forward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks* **5**(6): 989–993.

Haykin S (1999). *Neural Networks*: *A Comprehensive Foundation*. 2nd edition. Prentice-Hall, Englewood Cliffs, NJ, USA, pp. 156–255.

Hong H P and Goda K (2007). Orientation-dependent ground motion measure for seismic hazard assessment. *Bull Seismol Soc Am* **97**: 1 525–1 538.

Hong H P, Zhang Y and Goda K (2009). Effect of the spatial correlation on the estimated ground motion prediction equations. *Bull Seismol Soc Am* **99**(2A): 928–934.

Joyner W B and Boore D M (1993). Methods for regression analysis of strong-motion data. *Bull Seismol Soc Am* **83**(2): 469–487.

Kerh T and Chu D (2002). Neural networks approach and microtremor measurements in estimating peak ground acceleration due to strong motion. *Advances in Engineering Software* **33**(11–12): 733–742.

Lee S C and Han S W (2002). Neural-network-based models for generating artificial earthquakes and response spectra. *Computers and Structures* **80**(20-21): 1 627–1 638.

Lin C C J and Ghaboussi J (2001). Generating multiple spectrum compatible accelerograms using stochastic neural networks. *Earthquake Engineering and Structural Dynamics* **30**(7): 1 021–1 042.

Marquardt D W (1963). An algorithm for least squares estimation of non-linear parameters. Journal of the Society for *Industrial and Applied Mathematics* **11**(2): 431–41.

McGuire R K (2004). *Seismic Hazard and Risk Analysis*. Earthquake Engineering Research Institute, Oakland, CA, pp. 74–91.

Nguyen D and Widrow B (1990). Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. *Proceedings of the International Joint Conference on Neural Networks*. IEEE, San Diego, CA., **3**: 21–26.

Pacific Earthquake Engineering Research (PEER) Center (2006). Next Generation Attenuation database. http://peer.berkeley.edu/nga/index.html. (last accessed April 4th, 2006).

Press W H, Teukolsky S A, Vetterling A W T and Flannery B P (1992). *Numerical Recipes in C*: *The Art of Scientific Computing*. Cambridge University Press, New York. pp. 683–688.

Principe C, Euliano N R and Lefebvre W C (1999). *Neural and Adaptive Systems*: *Fundamentals through Simulation*. John Wiley and Sons, New York, NY, USA.

Rumelhart D E, Hinton G E and Williams R J (1986). *Learning Internal Representation by Error Backpropagation, in Parallel Distributed Processing*: *Explorations Microstructure of Cognition* (Vol. 1). MIT Press, Cambridge, Mass, USA, pp. 318–362.

Singh S K, Mena E, Castro R and Carmona C (1987). Empirical prediction of ground motion in Mexico City from coastal earthquakes. *Bull Seismol Soc Am* **77**: 1 862–1 867.

Wang H S (1993). Intelligent determination of peak ground motion parameters based on intensity assessed by artificial neural network. *Acta Seismologica Sinica* **6**(3): 705–712.

Youngs R R, Silva W J and Humphrey J R (1997). Strong ground motion attenuation relationships for subduction zone earthquakes. *Seism Res Lett* **68**: 58–72.