AN EARTHQUAKE DAMAGE MODEL USING NEURAL NETWORKS

by:
Gilbert L. MOLAS* and Fumio YAMAZAKI**

INTRODUCTION

Early estimation of damage due to earthquakes is an important concern in Japan. It is useful for city gas companies to decide whether to shut off the gas supplies following a large earthquake. If the damage is large, timely shutoff of the gas supply may prevent secondary disasters. However, if the gas supply was shut off unnecessarily, it might take time to restore the service and the inconvenience to the customers may be more serious. In an attempt to make an early but accurate estimate of damage to customers' buildings and pipelines, an extensive monitoring system of earthquake intensities was developed in Japan[1]. This system measures the peak ground acceleration (PGA) and the Spectrum Intensity (SI) at many points within a service area. The measured PGA and SI are transmitted by radio to the headquarters of the gas company where the damage estimation is conducted.

To estimate the damage from PGA and SI, however, is not an easy task. Obviously, if we specify the structures and input motion in terms of time history, sophisticated response analysis can be conducted. However, if we must estimate overall damage of many types of structures from the measured earthquake ground motion indices, a quick and robust method is necessary. The PGA is the most commonly used index to describe the severity of the earthquake ground motion. However, it is well known that a large PGA is not always followed by severe structural damage. Katayama et al.[2] demonstrated that the SI value has a better correlation with structural damage than PGA. Other indices of earthquake ground motion, e.g., peak ground velocity (PGV), peak ground displacement (PGD), duration of strong motion, and spectral characteristics of various definitions, can also be considered in such a damage estimation[3,4,5]. Ando et al.[6] demonstrated that PGA, PGV, and PGD are correlated to the damage of short-period, intermediate-period, and long-period structures, respectively.

Correlating the ground motion indices to the observed damage in a mathematical form is not easy because of the large uncertainties involved and the relationship must be highly nonlinear. A conventional way to construct such a relationship from observed data is to use multiple regression analysis. In such a case, a functional form must be assumed to relate input and output parameters. To avoid this, the use of neural networks is proposed in this paper for earthquake damage estimation.

Among several new techniques of artificial intelligence, neural networks or parallel distributed processing (PDP) has recently drawn considerable attention in various fields of science and technology. Along with the development of theories and computational algorithms[7,8], the technique has been applied to fields like automation, character recognition, electro-communication and noise filtering, image processing, industrial control problems, etc. Recently, it has been

* Graduate student, Institute of Industrial Science, University of Tokyo,
7-22-1 Roppongi, Minato-ku, Tokyo 106, Japan
** Associate Professor, ditto

applied to problems in earthquake engineering, e.g., active vibration control of structures[9], seismic hazard prediction[10].

Unlike expert systems that require human experts to formulate rules with which to arrive at a solution, neural networks need only examples of the input and output. Through a learning process, the neural network will attempt to find an internal state to represent the relationship between input and output parameters. The use of neural networks for earthquake damage estimation has several other advantages: once the network has been set up, damage estimation from new inputs is very fast and retraining the network for new data is relatively simple and can be done off-line. However, since the estimation is highly dependent on the learning data, we must prepare well-examined data sets.

TRAINING DATA

To construct a relationship between earthquake ground motion and structural damage, a data set comprising inputs (strong ground motion parameters) and outputs (damage) must be prepared. There are basically two methods for doing this: one is to collect actual earthquake records and damage data near the recording site; the other is to perform earthquake response analyses for given inputs and models and obtaining the resultant damage (outputs). The former is more convincing because it uses actual damage data. However, good recordings obtained near structural damage are few. With the latter, it is easier to prepare well-distributed data. Since it is not based on actual observations, however, much care should be taken in selecting structural models and input motions. The former was used by the authors and reported elsewhere[11]. The overall damage was given in three classifications (i.e., negligible, moderate, and severe) based on the observed damage to buildings and pipelines. The classification was based on extensive literature survey and some site investigations. It was found that although the resulting neural network can give good estimates for negligible and severe damage, the data is not enough to estimate moderate damage well. For this paper, the latter method is used. However, since recorded ground motions are not well distributed within the expected range of values, we use simulated ground motion for the analysis.

Simulation of Strong Ground Motion

The Kanai-Tajimi (K-T) power spectrum is used to generate stationary time series that are then multiplied by a trapezoidal envelope function to represent the duration of motion and the rise and decay of the ground motion. In this study, the rise and decay times of the ground motion are assumed to be constant at 2.5 s. The Kanai-Tajimi power spectrum is defined as

$$S(\omega) = S_0 \cdot \frac{1 + 4h_g^2 \omega^2 / \omega_g^2}{(1-\omega^2 / \omega_g^2)^2 + 4h_g^2 \omega^2 / \alpha} \qquad (1)$$

where ω is the circular frequency under consideration, S_o is the intensity, ω_g is the K-T frequency, and h_g is the K-T damping.

Lai[12] studied the statistical characteristics (e.g., mean, standard variation) of the K-T parameters based on actual earthquakes. Although he determined probability density functions

Table 1. Parameters for earthquake generation

Parameter	Lower Limit	Upper Limit
So	1.0 cm^2/s^3	500.0 cm^2/s^3
wg	4.0 rad/s	40.0 rad/s
	(0.64 Hz)	(6.34 Hz)
hg	0.15	0.60
total time	7.5 s	20.0 s

for the K-T parameters, it is desired to have a uniformly distributed set of parameters to have a well-distributed data set. Hence, the values of the parameters used to generate the artificial earthquake motions are randomly selected from a range of typical values (Table 1) assuming a uniform probability distribution. In this study, 500 artificial ground acceleration time series are used. Figure 1 shows the distribution of the maximum acceleration and SI for the set of simulated ground motions.

As the simplest indices of ground motion severity, the PGA, PGV, PGD, and SI of the input ground motion are considered. Note that in this study, the SI value is defined as the average velocity response spectrum of 20% damped single-degree-of-freedom systems with natural period between 0.1 s to 2.5 s as[2]

$$SI = \frac{1}{2.4} \int_{0.1}^{2.5} S_V(T,h=0.2)\, dT \tag{2}$$

In addition, the root square, R_s, of the acceleration defined as

$$R_s = \sqrt{\int_0^T a^2(t)\, dt} \tag{3}$$

where $a(t)$ is the ground acceleration, is used to account for the total power of the ground motion. The time duration of the ground motion, T_d, defined by Trifunac and Brady[13] as the time where the middle 90 percent of the total power is realized, is also used.

Structure / Damage Models

To estimate the damage of structures due to strong ground motion, the nonlinear response of single-degree-of-freedom (SDOF) models using the Newmark-β method is used. Two SDOF models that represent two types of wooden framed structures commonly found in Japan are used. The first model (Wooden 1) represents ordinary wooden framed houses with two stories and a fundamental period of T=0.55 s. The second (Wooden 2) represents fire-resisting wooden

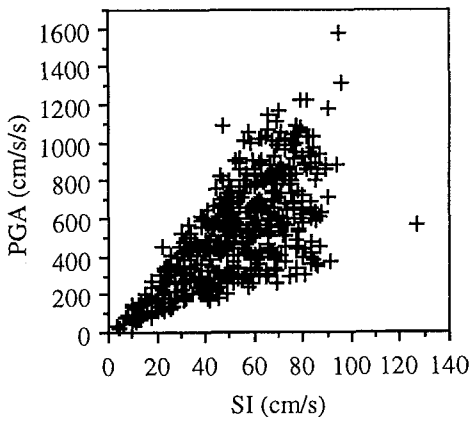


Figure 1. Distribution of the peak ground acceleration and spectral intensity for the simulated earthquake ground motions

framed houses with two stories and a fundamental period of T=0.35 s. The models have bilinear stiffness with the secondary stiffness taken as 20% of the initial stiffness. The damping ratio is taken to be 0.05 and the restoring force at yielding is[14]

$$Q_y = mg \cdot C_y \qquad \text{where } C_y = 0.25 / \sqrt{T} \tag{4}$$

and m is the mass (taken as unity), g is the acceleration due to gravity (=980 cm/s^2), and T is the fundamental period (s). The damage to the structure is then given in terms of the ductility factor, μ, defined as

$$\mu = \frac{U_{max}}{U_y} \tag{5}$$

where U_{max} is the maximum displacement by a step-by-step bi-linear analysis and U_y is the yield displacement.

It should be noted that the structure/damage models and the strong ground motion model employed here are rather simple. However, the main purpose of this paper is to demonstrate the use of neural networks for the quick estimation of damage, and more sophisticated models can be introduced in a future study using the same procedure.

NEURAL NETWORK MODEL

A neural network is a collection of parallel processors connected in the form of a directed graph. A network consists of neurons or Processing Elements (PEs) which are arranged in layers. The neural network structure used is a three-layered feed-forward neural network with full connectivity and bias (Figure 2). The bottom layer called the input layer holds the input vector and has one PE for each variable in the input vector plus an optional bias. The top layer called the output layer holds the output values of the network. In between the input and output layers, there can be one or more hidden layers with different number of PEs. We use one hidden layer with four PEs and bias. It widely known[15,16] that one hidden layer is generally sufficient for back-propagation networks. A single hidden layer network is also easier to train and gives excellent results. We found that networks with more than four PEs in the hidden layer does not really give significant improvement but takes more time to train, at least, in the case of this problem.

Input data is fed to the input layer and processing is done layer-by-layer up to the output layer.

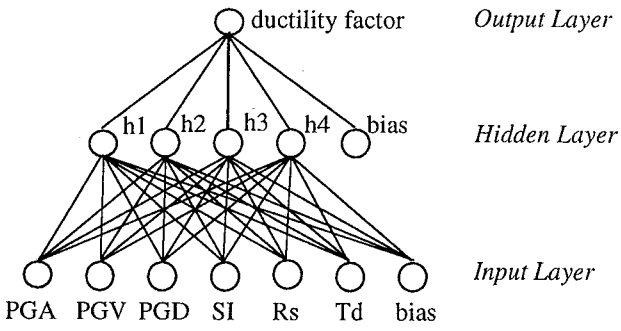


Figure 2. Neural network structure for this study

The output values of the hidden layer and output layer PEs can be expressed as[15]

$$out_j^h = f^h\left(\sum_{i=1}^{N} \omega_{ji}^h x_i + \theta_j^h\right) \quad \text{and} \quad out_k^o = f^o\left(\sum_{j=1}^{M} \omega_{kj}^o out_j^h + \theta_k^o\right) \tag{6}$$

respectively, where ω_{ji} is the connection weight of the jth PE from the ith PE of the input layer, x_i is the ith scaled input, θ_j is the bias term for the jth PE and f is the transfer function between two layers. The superscripts define the variables for the outer layer and the hidden layer. Although neural networks are based on parallel processors, it is easy to simulate the computation process using sequential computers. Given randomized initial values of the weights and biases, the network output can be computed for a given input vector. The connection weights are then updated to decrease the difference between the network output and the desired output. The bias term is similarly updated by treating it like a weight with unity as its input value. In this study, the weights are updated after one complete pass of the training data set. The training is stopped after the error becomes less than a given value or has become stable.

The performance of two variations of the back-propagation algorithm; namely 1) the Normalized Cumulative Delta Rule[17] (*ncdr*) and 2) the Extended Delta-Bar-Delta Rule[18] (*edbd*); and a random search algorithm called the Directed Random Search[19] (*drs*) is presented. To examine the performance of the transfer functions, two functions, namely the sigmoid and hyperbolic tangent (*tanh*) functions, are compared. The input and output values are scaled based on the minimum and maximum values of the training data according to

$$\tilde{x} = l_L + (l_H - l_L) * \frac{x - x_{min}}{x_{max} - x_{min}} \tag{7}$$

where $\tilde{x}$ is the scaled value of x, l_L and l_H are the lower and upper limits of the scaled value, respectively, and x_{min} and x_{max} are the minimum and maximum values, respectively of parameter x in the training data set. The l_L and l_H of input values are -1.0 and 1.0, respectively. For output values, l_L is -0.8 in the case of the *tanh* transfer function or 0.2 in the case of the sigmoid transfer function, while l_H is 0.8 for both transfer functions. Scaling of the data values is needed to prevent the saturation of the transfer functions and to normalize the influence of input parameters with different units. The x_{min} and x_{max} values for the input and output values are given in Table 2. The computations for the three algorithms are done using the NeuralWorks Professional II/Plus Simulator[16]. Figure 3 shows the learning convergence for selected epoch counts (the number of

Table 2. Minimum and maximum values used in scaling

Parameter	xmin	xmax
PGA (cm/s2)	22.49	1573.78
PGV (cm/s)	3.09	111.12
PGD (cm)	0.66	43.40
SI (cm/s)	3.52	126.99
Rs (cm/s3/2)	23.58	1631.02
Td (s)	3.09	15.15
m (Wooden 1)	0.16	7.15
m (Wooden 2)	0.11	7.30

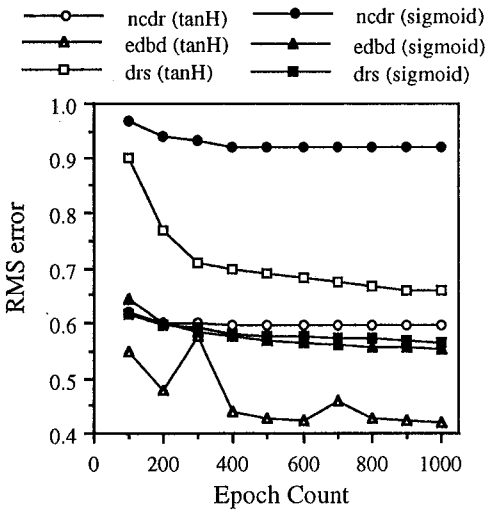


Figure 3. Performance curves of the three learning algorithms using the sigmoid and hyperbolic tangent transfer functions (for structure model Wooden 1)

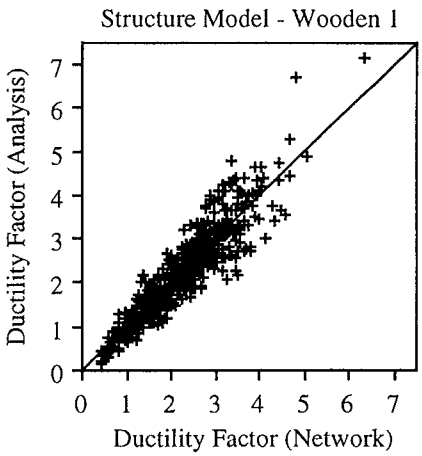


Figure 4. Comparison of ductility factor based on analysis and neural network for the training data set

weight updates) in terms of the root-mean-square error defined as

$$E_{RMS} = \sqrt{\frac{1}{N}\left(\sum_{n=1}^{N}(des_n - out_n)^2\right)} \tag{7}$$

where des_n and out_n are the desired and network output of the output layer PE for the nth input vector, respectively. N is the number of training data. It can be seen that for the back-propagation algorithms the *tanh* transfer function gives better performance and that the *edbd* algorithm using the *tanh* transfer function gives the best performance. For subsequent computations, the *edbd* algorithm with *tanh* transfer function is used. After training, the output of the neural network for the whole training data is compared to the desired output based on the bi-linear analysis in Figure 4. The results show good correlation for both structure models in the recall tests.

RESULTS

Sensitivity Analysis

Although neural networks can find a relationship between the input and output values internally, it is not always easy to interpret the resulting weight state. Table 3 shows the resulting weights and biases after training for model Wooden 1. No general trends regarding the weight states can be deduced by simple inspection because all the weights and biases are interrelated. Thus, the effect of one input parameter to the output is difficult to analyze. Alternatively, it is possible to compute the sensitivity of the output value with respect to one of its input by taking the partial derivative. From Equation 6, the partial derivative of an output PE, out_k^o, with respect to an input parameter, x_n, is then

Table 3. Weights and bias values for model Wooden 1

ω^o_{kj} \ j / k	h1	h2	h3	h4	Bias
Out	-0.9793	-1.6648	0.4552	-3.0525	-0.2723

ω^h_{ji} \ i / j	PGA	PGV	PGD	SI	Rs	Td	Bias
h1	0.6847	-0.2302	-0.6975	0.7066	0.0630	-0.1105	0.6531
h2	1.8667	-0.2285	-0.3556	-0.8487	0.9808	-0.0761	0.8459
h3	0.3843	-0.5458	-0.1407	1.7752	-0.4337	0.4247	-1.2121
h4	-1.2547	0.1191	0.4160	0.0725	-0.5285	0.0789	-0.8422

$$\frac{\partial}{\partial x_n}\left(out^o_k\right) = f'^o\left(\sum_{j=1}^{M}\omega^o_{kj}\cdot f^h\left(\sum_{i=1}^{N}\omega^h_{ji}\,x_i+\theta^h_j\right)+\theta^o_k\right)\cdot \sum_{j=1}^{M}\left(\omega^o_{kj}\cdot f'^h\left(\sum_{i=1}^{N}\omega^h_{ji}\,x_i+\theta^h_j\right)\cdot\omega^h_{jn}\right) \tag{8}$$

It can be seen from the above equation that the partial derivative depends not only on the weights and biases but also on the current values of the input variables, x_i's Thus, it is difficult to generalize on the trend of the output value with respect to a change in a single input value. However, the distribution of the partial derivatives for the entire training set can be used to qualitatively describe the sensitivity of the output value. Figure 5 shows the histograms of the partial derivatives for structure model Wooden 1. It can be seen from the scatter about the zero point that the output is more sensitive to the PGA, SI, and R_s and least sensitive to the time duration of motion, T_d. Figure 6 shows the plot of the ductility factor, m, with respect to each of the input parameters. From this figure, it can be seen that the ductility factor is uncorrelated to T_d for this training data set. This is a direct consequence of the assumed uniform distribution of the total time of simulation together with the Kanai-Tajimi parameter, S_0. Low intensity shaking may have large T_d and high intensity shaking may have small T_d. The same trend can be observed for structure model Wooden 2.

Input parameter selection

The previous section has identified the input parameters that have the most effect to the damage. It is then interesting to see the effect of using a reduced number of input parameters to the performance of the network. Of particular interest is the estimation using just PGA and SI since these indices are measured directly by a new type of seismometer of a Japanese gas company[1]. The summary of results is shown in Table 4. The trend is that more input parameters will generally give better results. However, if the input parameters that are most influential to the output are used (i.e., PGA, SI, and R_s), the result is comparable to the one using all input parameters. The estimation using just PGA and SI is also comparable to the best estimation. It should also be noted that the appropriate parameters for damage estimation are structure

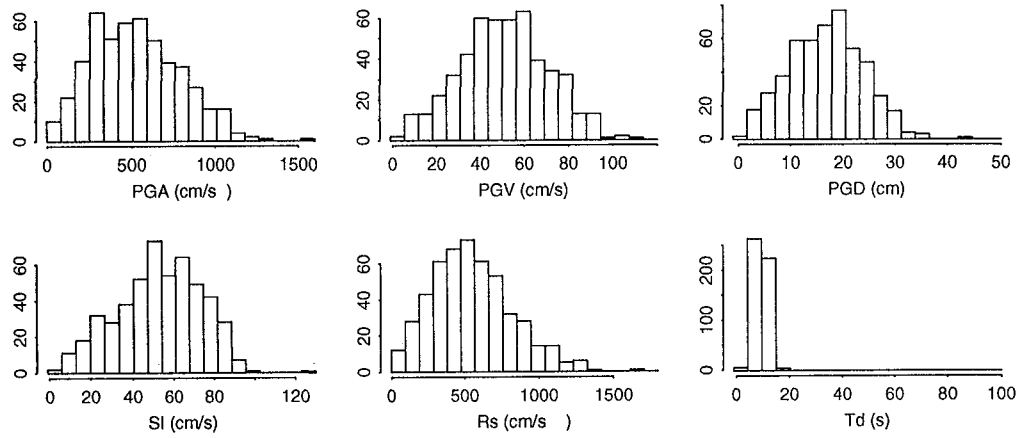


Figure 5. Histograms of the partial derivatives of the scaled output with respect to the scaled inputs for structure model Wooden 1

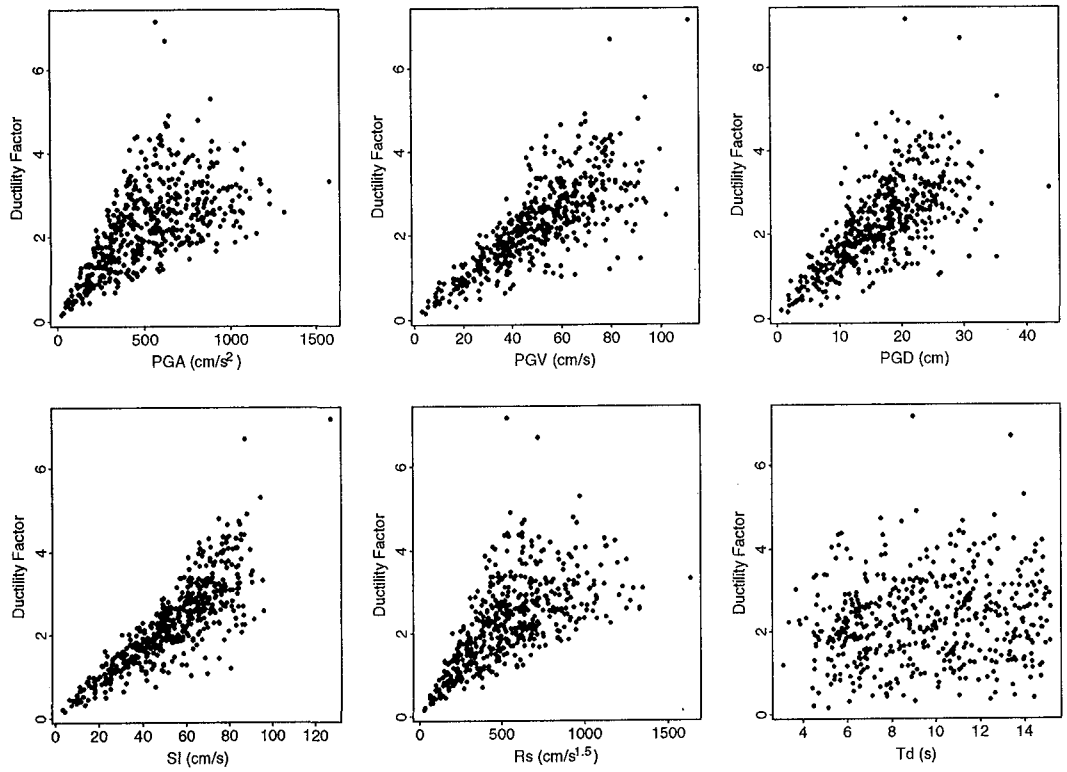


Figure 6. Plot of input parameter vs. ductility factor for structure model Wooden 1

Table 4. Root-mean-square error and correlation between analysis and neural network after training (Simulated earthquakes, No. of data = 500)

Input Parameters	Wooden 1		Wooden2	
	RMS error	Correlation	RMS error	Correlation
PGA, PGV, PGD, SI, Rs & Td	0.406	0.913	0.454	0.924
PGA only	0.747	0.665	0.703	0.809
SI only	0.546	0.837	0.833	0.717
PGA & SI	0.434	0.900	0.493	0.910
PGA, SI & Rs	0.424	0.904	0.479	0.915
PGA, SI & Td	0.429	0.902	0.482	0.914
PGA, PGV, PGD, SI & Rs	0.413	0.910	0.478	0.915

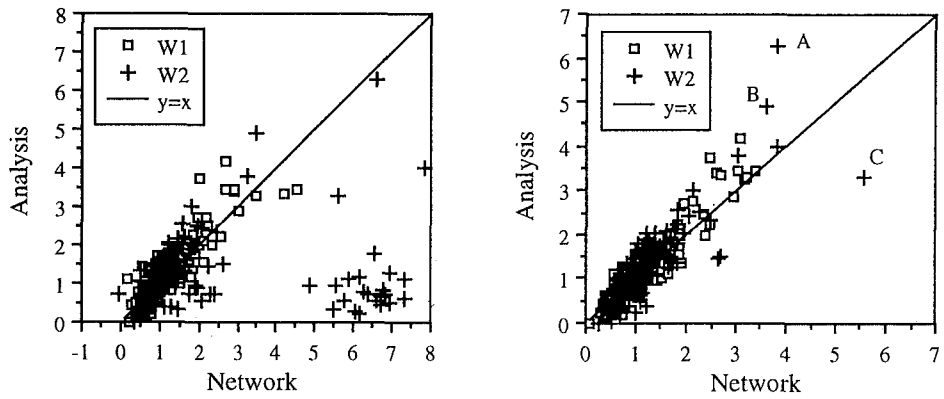
Table 5. Summary of earthquake events used in this study

Earthquake Event	No. of records (components)	
Niigata (1964)	1	(2)
Matsushiro (1965-66)	23	(46)
Tokachi-Oki (1968)	3	(6)
Miyagiken-Oki (1978)	4	(8)
Nihonkai-Chubu (1983)	2	(4)
Chibaken-Toho-Oki (1987)	9	(18)
Izu Pen. Toho-Oki (1989)	4	(8)
Kushiro-Oki (1993)	5	(10)
Noto Pen. Oki (1993)	1	(2)
Imperial Valley (1940)	1	(2)
San Fernando (1971)	5	(8)
Mexico (1985)	2	(4)
Loma Prieta (1989)	19	(38)
Total	79	(156)

dependent. For Wooden 1 (T=0.55 s), "SI only" shows much better performance than "PGA only". However, for Wooden 2 (T=0.35 s), the PGA is a better index than SI. This observation is the same as the one by Ando et al.[6]

Estimation using actual earthquakes

To test the performance of the trained networks in the case of actual strong ground motion, the time history of actual earthquakes used by Yamazaki et al.[11] are considered. Since these time histories were not used to train the neural networks, the behavior of the estimates would show how well the trained network will perform in an actual implementation. Table 5 shows a summary of the earthquakes and the number of records for each earthquake used. In the case of Wooden 1, the data cases the network outputs and analysis outputs for the 156 time histories have good correlation. However, there are several gross over-estimations for the case of Wooden 2 as shown in Figure 7a. This behavior is seen as an instability of the network due to the use of an input not



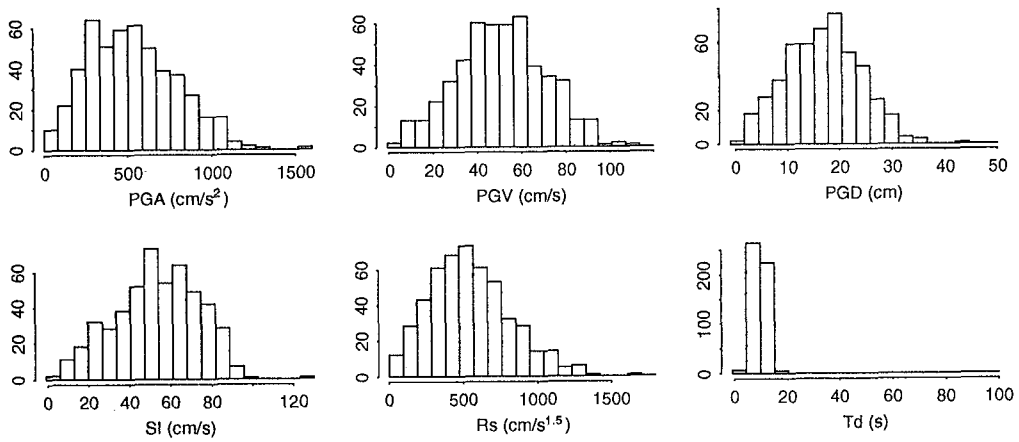
(a) Using PGA, PGV, PGD, SI, R_s & T_d

(b) Using PGA and SI

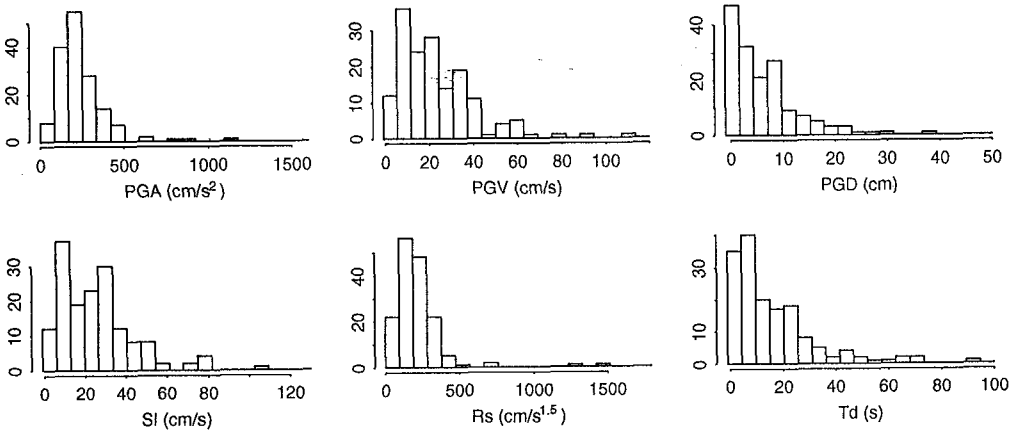
Figure 7. Performance of two trained neural networks if recorded earthquake ground motions (156 components) are used as input

within the range of the training data. By comparing the range of input parameters for the simulated and actual earthquake records (Figure 8), it was observed that the T_d for the actual records greatly exceeded the range of the simulated earthquakes. If T_d is not used as an input, the instability is eliminated. Based on the sensitivity analysis, the effect of not using T_d in the estimation will be small. Another way to eliminate the instability is to simulate new strong motions by increasing the range of the total time in Table 1 and retrain the network. But the difference in the estimation compared to the case of not using T_d as an input will be small. Table 6 shows the performance of the neural networks in Table 4 for the actual earthquake ground motions. Note that the network that uses PGA and SI only gave the best performance.

Also of particular interest are three data points marked "A", "B", and "C" in Figure 7b. These points have large errors that cannot be attributed to instability. "A" corresponds to the acceleration time history of the East-West component of the 1993 Kushiro-Oki earthquake in Japan recorded at JMA Kushiro Station (PGA=920 cm/s^2, SI=78 cm/s). "B" corresponds to the S74°W component of the 1968 San Fernando earthquake recorded at Pacoima Dam (PGA=1055 cm/s^2, SI=74cm/s) while "C" correspond to the S16°E component (PGA=1148 cm/s^2, SI=105 cm/s). These ground motions are very strong and are in the range where the training data are sparse (Figure 6). Hence, when we apply a trained network to new input data, we must examine whether the data exist within the range of the training data. It must be emphasized that neural networks are learning data dependent and that to prepare a good learning data set is the most important issue in the use of neural networks.



(a) Simulated ground motions (Number of data = 500)



(b) Recorded ground motions (Number of data = 156)

Figure 8. Histograms of input parameters of a) simulated ground motions used in training; and b) recorded ground motions

Table 6. Root-mean-square error and correlation between analysis and output of trained networks in Table 4 using actual earthquake ground motion (No. of data = 156)

Input Parameters	Wooden 1		Wooden2	
	RMS error	Correlation	RMS error	Correlation
PGA, PGV, PGD, SI, Rs & Td	0.373	0.883	2.090	0.175
PGA only	0.617	0.691	0.503	0.792
SI only	0.378	0.879	0.608	0.840
PGA & SI	0.320	0.920	0.446	0.840
PGA, SI & Rs	0.332	0.907	0.459	0.830
PGA, SI & Td	1.229	0.400	0.755	0.638
PGA, PGV, PGD, SI & Rs	0.375	0.882	0.480	0.814

CONCLUSIONS

The use of neural networks to predict damage from simple ground motion indices is demonstrated. Since the strong motion parameters of recorded accelerograms are not well distributed, simulated ground motions are used. The training data is generated by computing ground motion indices of the simulated earthquakes. The damage is given in terms of the ductility factor computed from a step-by-step bi-linear analysis of two structural models representing wooden houses commonly found in Japan. Neural networks are found to be useful in finding a relationship between ground motion indices and the corresponding damage. The neural network acts as a transfer function with the ground motion indices as input and damage as output. The trained neural network is tested by using actual earthquake ground motions to compute the indices and ductility factor. It is found that the trained network performs well. However, careful attention must be paid to the range of values of the input vector, as this could lead to instability in the predictions. New input vectors used for estimation must be within the range of the training data vectors. The input parameters that have greater influence to the output can be identified qualitatively after training by taking the partial derivative of the output with respect to an input variable for the entire training data set. The PGA, SI, and root square are found to be the most influential to the damage. The damage estimation using PGA and SI is comparable to the best estimation when using the training data and gave the best estimation when using actual earthquake ground motion. Although the analysis is limited to two structure models, other structure types can be similarly analyzed and the idea may be conveniently used in the damage estimation based on earthquake monitoring.

References:

1. H. Nakane, E. Kodama, and I. Yokoi, 'New earthquake-induced ground motion severity sensing apparatus for reliable system shutdown', *Proc. 10th World Conf. on Earthq. Eng.* **9** 5559-5562 (1992).
2. T. Katayama, N. Sato, and K. Saito, 'SI-sensor for the identification of destructive earthquake ground motion', *Proc. 9th World Conf. on Earthq. Eng.* **7**, 667-672 (1988).

3. R. Araya and G.R. Saragoni, 'Earthquake accelerogram destructiveness potential factor,' *Proc. 8th World Conf. on Earthq. Eng.* **2**, 835-842 (1984).

4. Y.J. Park, A.H-S Ang, and Y.W. Wen, 'Seismic damage analysis of reinforced concrete buildings,' Proc. ASCE Jour. Structural Division **111**, 740-757 (1985).

5. T.J. Zhu, A.C. Heidebrecht, and W.K. Tso, 'Effect of peak ground acceleration to velocity ratio on ductility demand of inelastic systems,' Earthquake Engineering and Structural Dynamics **16**, 63-79 (1988)

6. Y. Ando, F. Yamazaki, and T. Katayama, 'Damage estimation of structures based on indices of earthquake ground motion', *Proc. 8th Japan Earthq. Eng: Symposium* **1**, 715-720 (1990) (in Japanese).

7. I. Aleksander and H. Morton, *An introduction to neural computing,* Chapman & Hall, 1990.

8. P. Antognetti and V. Milutinovic (Eds.), *Neural Networks: Concepts, applications, and implementations Volume IV*, Prentice Hall, 1991.

9. Y. Nekomoto, K. Fujita, and M. Tanaka, 'Simulation on the active vibration control of structures by neural networks', *Proc. Symposium on Application of Fuzzy and Neural Networks, JSME:* 93-96 (1991) (in Japanese).

10. F.S. Wong, A.T.Y. Tung, and W. Dong, 'Seismic hazard prediction using neural nets', Proc. 10th World Conf. on Earthq. Eng. 1, 339-343 (1992).

11. F. Yamazaki, G.L. Molas, and M. Fatima, 'Use of Neural Networks for Earthquake Damage Estimation', Proc. 6th ICOSSAR (1993) (in press).

12. S.P. Lai, 'Statistical characterization of strong ground motions using power spectral density function', BSSA 72:1, 259-274 (1982)

13. M.D. Trifunac and A.G. Brady, 'A study of the duration of strong earthquake ground motion' BSSA 65, 581-626 (1975).

14. Mitsubishi Research Institute, Seismic Damage Assessment of Tokyo, Research Report of Building Sub-committee, 1989 (in Japanese).

15. J.A. Freeman and D.M. Skapura. Neural Networks (Algorithms, Applications, and Programming Techniques). Addison-Wesley, 1991.

16. NeuralWare, Inc., Neural Computing: NeuralWorks Professional II/PLUS and NeuralWorks Explorer, 1991.

17. J.L. McClelland and D.E. Rumelhart, Explorations in Parallel Distributed Processing (A handbook of models, programs and exercises), MIT Press, 1988.

18. A.A. Minai and R.D. Williams, 'Acceleration of back-propagation through learning rate and momentum adaptation', Intn'l Joint Conf. on Neural Networks 1, 676-679 (1990).

19. N. Baba, 'A new approach for finding the global minimum of error function of neural networks', Neural Networks 2, 246-253 (1989).