eJSE International

Electronic Journal of Structural Engineering 12(1) 2012

Artificial Neural Network Application for Predicting Seismic Damage Index of Buildings in Malaysia

Azlan Adnan & Patrick Liq Yee Tiong

Engineering Seismology & Earthquake Engineering Research (e-SEER), Universiti Teknologi Malaysia

Rozaina Ismail

Faculty of Civil Engineering, Universiti Teknologi MARA, Malaysia

Siti Mariyan Shamsuddin

Department of Computer Graphics and Multimedia, Universiti Teknologi Malaysia, Malaysia

Email: tiong.patrick@ymail.com

ABSTRACT: An effective, convenient and reliable intelligent seismic evaluation system for buildings in Malaysia has been developed in this study by using Back-Propagation Artificial Neural Network (ANN) algorithm. A total of forty one buildings with 164 sets of input data spreading throughout Peninsular and East Malaysia were chosen and analyzed using IDARC-2D finite element software under seismic loading at peak ground accelerations of 0.05g, 0.10g, 0.15g and 0.20g respectively. Non-linear dynamic analysis was performed in order to obtain the damage index of each building. The ANN algorithm comprising 15 hidden neurons with 1 hidden layer outperformed other combinations in predicting the damage index of buildings with accuracy statistical value of 93% in testing phase as well as 75% in validation stage. From the results, the ANN system is suitable to be used for predicting the seismic behaviour of their buildings at any given time.

KEYWORDS: Seismic performance of buildings, Artificial Neural Network, damage index of building

1 INTRODUCTION

The occurrence of earthquake itself is not catastrophic. Disaster from earthquake happens when the shaking of ground causes severe damage onto structures built by mankind. Thus, to prevent tragedies from happening due to earthquake, the effectiveness of seismic structural evaluation technique is essential. However, the key question arises. How strong an earthquake is classified as strong? At what intensity of shaking will a particular building need to be evacuated? For earthquake prone countries such as Indonesia where small magnitude of ground shakes occur frequently, it might not be practical for evacuation of buildings whenever tremors are felt. To make matter worse, a similar magnitude of tremor might be causing multiple types of effects onto different types of structures depending on the natural period of each building. Therefore, an intelligent seismic evaluation system was innovated in this study by incorporating Artificial Neural Network (ANN) algorithm in predicting damage index of a particular given building at any given time.

2 LITERATURE REVIEW

2.1 *Artificial Neural Network (ANN)*

Artificial Neural Network (ANN) is typically referring to mathematical algorithm model imitating behaviours of biological neuron system. Generally, neural networks comprise neurons which are expressed mathematically in computer programming languages to behave like human brains. The algorithm usually consists of three basic layers; input layer, hidden layer and output layer respectively (Figure 1). Numerous researchers (Wu *et al*, 1992; Masters, 1993; Tang *et al*, 1993; Tsou and Shen, 1993; Zhao *et al*, 1998) had investigated the capabilities of utilizing ANN in the field of structural engineering. Most of them agreed that ANN was capable to solve at least as well as a traditional method for any given task.

Very much alike to human brains, the artificial neural networks have the capability to learn apply knowledge from past experiences in solving new problems in new environment (Abdullah *et al*, 2000). This learning is provided to the neural network through training process where an amount of data is fed into the neural network in order to obtain the optimum weight values within the algorithm. There are typically two types of learning methods available, namely the supervised and unsupervised method. The supervised method being more common is adapted in this study.

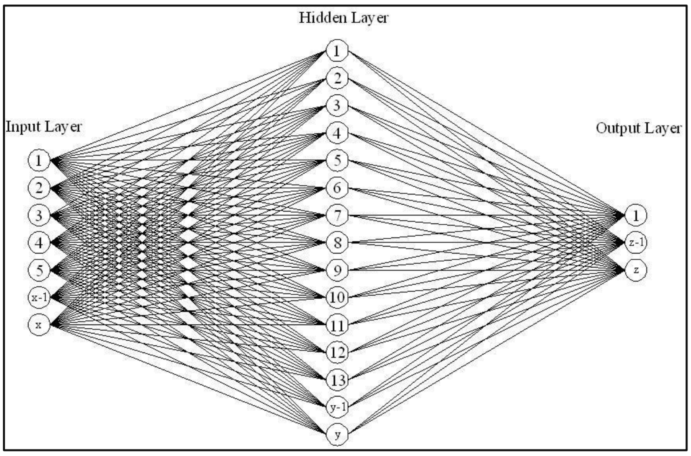

Figure 1. Typical Structure of ANN

2.2 *Damage Index (DI)*

Most of the existing building performance level indicators are expressed in qualitative forms rather than in numerical values. Taking for example, the Structural Engineers Association of California, SEAOC has identified four level of performance indicators; namely Fully Operational (FO), Operational (O), Life Safe (LS) and Near Collapse (NC) in their Vision 2000 guidelines (SEAOC, 1995). Besides that, the National Earthquake Hazard Reduction Program (NEHRP) Guidelines have proposed similar performance indicators as in Vision 2000.

However IDARC-2D, the chosen analysis software in this study incorporates the Park & Ang damage model (Park *et al.*, 1984). According to the model, Park & Ang damage index ($DI_{P\&A}$) for a particular structural element is determined by:

$$DI_{\mathrm{P\&A}} = \frac{\delta m}{\delta u} + \frac{\beta}{\delta_{\mathrm{u}} \mathrm{P_y}} \int dE_{\mathrm{h}} \qquad (1)$$

where δ_m = maximum experienced deformation; δ_u = ultimate deformation of element; P_y = yield strength of element; $\int dE_{\mathrm{h}}$ = hysteretic energy absorbed by element during response history; and β = model constant parameter.

The Park & Ang damage model has been calibrated with nine separate observations of damaged reinforced concrete buildings (Park *et al.*, 1986). The calibrated damage indexes which represent the level of physical deterioration of those associated buildings are listed in Table 1.

Table 1. Calibrated damage indexes of Park & Ang damage model (Park *et al.*, 1986)

Damage Level	*Physical Condition*	*Damage Index*
Collapse	At least partial collapse	> 1.0
Severe	Widespread concrete crushing or reinforcement buckling	0.4 – 1.0
Moderate	Visible cracks or concrete spalling	< 0.4

Based on the damage index of an individual structural element according to (1) is inadequate to categorize the global damage of the whole building. Therefore, Kunnath *et al*., (1992) has made modification to the previous damage model due to complication in incorporating the relative storey movements with localized element plastic hinge deformation. This modified damage model is shown in (2).

$$DI = \frac{\theta_m - \theta_r}{\theta_u - \theta_r} + \frac{\beta}{M_y \theta_u} E_h \qquad (2)$$

Where DI = modified damage index; θ_m = maximum rotational response; θ_u= ultimate sectional rotational capacity; = θ_r unloading recovered rotation; M_y = yield moment; and E_h = dissipated energy within the section

2.2 *Current Practice of Seismic Inspection*

The most widely adopted seismic inspection currently available is the so-called rapid screening procedure or in short, the RSP survey. This method comprises two different sets of protocols or procedures for inspectors to rapidly evaluate the seismic risk of a particular building of interest merely via visual inspection. The first method is known as ATC-21survey (ATC-21, 1988). This method is designed to serve as preliminary tools in assessing the building's capability to resist seismic threats just by judging from its external appearance. Using the predetermined checklist, inspectors will then decide whether the building is hazardous based on the score sheet.

Those buildings deemed unsatisfactory will then have to undergo ATC-22 evaluation (ATC-22, 1989). In this checklist, the structural integrity as well as non-structural implications is taken into consideration. Nevertheless, the assessment is still concentrating on qualitative evaluation based on the score sheet contained on the checklist.

2.3 *Research Significance*

As discussed previously, the currently available seismic inspection methods (ATC-21 and ATC-22) are unable to generate quantitative damage indexes. There exists a gap of knowledge to implement the quantitative Park & Ang damage model in assessing real structures, without carrying out thorough structural analysis in advance. The main objective of this study was to investigate the feasibility in adopting Artificial Neural Network (ANN) algorithm in predicting seismic damage of reinforced concrete frame buildings under damaging earthquake forces. Up to the authors' knowledge, there is no extensive research information available regarding such subject matter up-to-date. Therefore, this study also examined the types of parameters to be fed into the ANN system as input in order to obtain higher accuracy of building damage predictions.

3 METHODOLOGY

The main research methodology adopted in this study could be categorized into two different main phases. The first phase will be the nonlinear time-history finite element analysis of the building samples by IDARC-2D to obtain the damage index for each particular building.

In performing the nonlinear response history analyses using IDARC-2D, the 1940 El-Centro ground motion data was scaled to four levels of peak ground acceleration (PGA) value; 0.05g, 0.10g, 0.15g and 0.20g. As the nature of this study was to determine the feasibility in applying the ANN model in damage prediction, the finite element modeling was just serving the purpose of generating input data to be fed into the ANN algorithm. Hence, the variability of frequency content of the time history was not the main context in this phase of study. Nevertheless, the predominant period of the scaled El-Centro time history was considerably covering a vast range from 0.01 to 3.0s (Figure 2).

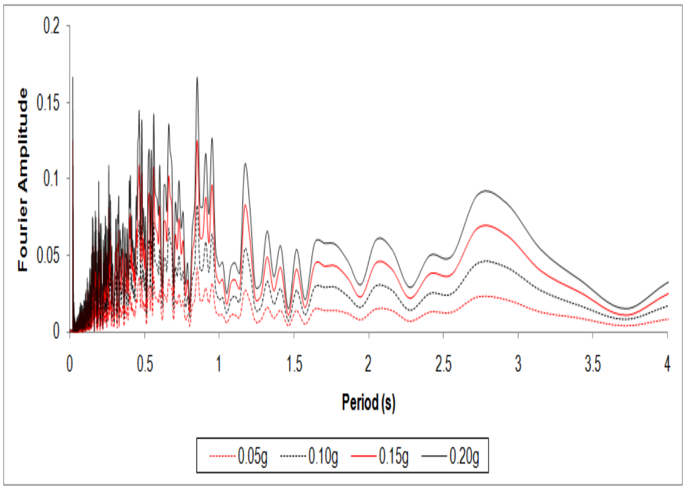

Figure 2. Predominant period of El-Centro response history

The second phase comprised the development of Back-Propagation (BP) Artificial Neural Network (ANN) algorithm for prediction of damage index for buildings at any given seismic accelerations. This paper focuses on the latter phase of the study that is regarding the data characteristics and classifications for the ANN process. Besides that, the vitality of each identified parameters towards the damage index prediction was determined by mathematical models (mean square method and linear correlation coefficient analysis). The thorough finite element modeling and analysis of the building samples were being reported elsewhere (Ismail, 2008).

In order to generate training data as input for the artificial neural network system, structural data of a total of twenty seven buildings scattered around Peninsular and East Malaysia were collected, and modeled for nonlinear response history analyses. These twenty eight buildings would generate 112 sets of total samples. Due to national security and privacy issues, the names of these buildings were kept anonymous throughout this paper. Nevertheless, some basic information particularly dimensional detail and dynamic property of these buildings are listed in Table 2. The initial natural period of these buildings, ranging from 0.22 to 1.86s fell within the predominant period of the scaled 1940 El Centro ground motion data.

Table 2. Brief information of the sample buildings for training data

Building	*Length* (m)	*Height* (m)	*Bay* (nos.)	*Age* (years)	*Period* (s)
1	6.1	48.5	1	33	1.03
2	72.6	16.6	11	13	0.84
3	24.0	19.2	3	13	0.46
4	19.2	25.7	3	27	1.86
5	36.0	7.4	6	9	0.29
6	11.0	7.5	2	23	0.36
7	42.0	13.8	14	6	0.50
8	10.0	13.8	3	4	0.36
9	20.1	39.9	3	28	0.83
10	6.1	34.9	1	23	0.80
11	75.6	20.0	9	9	0.83
12	6.1	34.9	1	28	0.81
13	11.0	17.0	3	26	0.47
14	9.0	36.2	3	37	0.94
15	18.0	15.2	4	5	0.23
16	20.1	14.3	3	5	0.22
17	9.0	10.6	3	4	0.40
18	9.0	17.0	3	23	0.22
19	48.0	7.4	8	11	0.39
20	72.6	21.6	11	58	0.40
21	24.0	23.2	3	22	0.48
22	16.0	51.0	2	16	1.01
23	7.5	19.0	2	22	0.48
24	12.2	39.0	2	25	0.97
25	48.0	23.0	6	45	0.72
26	8.0	39.0	1	46	0.87
27	19.5	73.0	3	16	1.83
28	24.0	64.2	3	31	1.64

3.1 *Parameters affecting building damage index*

This study is trying to conduct non-linear correlation between seven identified factors or parameters that might contribute to different seismic damage index for different buildings. Five of the parameters belong to the characteristics of buildings; age, number of bay, height, length and natural period of each building respectively. Meanwhile, the remaining two parameters are dependent on the seismic force natures; seismic zone and ground acceleration. These seven parameters were identified and established through the non-linear IDARC-2D analysis of all the building samples. The significance of each parameter in influencing the overall damage index predicted was tested by performing both the mean square error (MSE) method and linear correlation coefficient analysis respectively, which will be presented in detail in relevant section later.

3.2 *Artificial Neural Network (ANN) Development*

The algorithm of the ANN system was developed with Microsoft C++ programming language. The main of advantage of in-house developed software is that it allows for customization and better tuning capability. The developed ANN algorithm consisted of mainly 4 main phases (Figure 3).

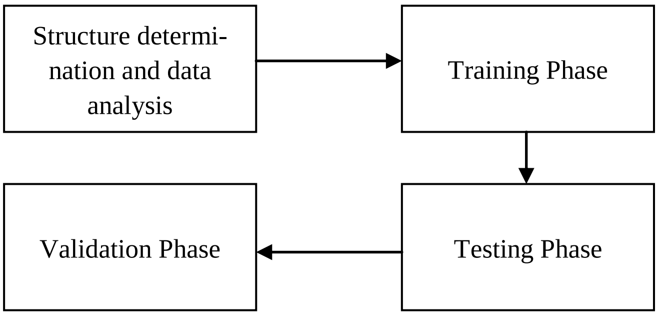

Figure 3. Main phases in ANN algorithm

The pioneer stage involved determination of output required from the ANN system. During this stage, the input parameters and numbers of hidden neurons were investigated. Sample data from IDARC-2D were analyzed, classified and distributed. The distributed data were then fed into the ANN algorithm for testing phase. In training algorithm, the network would perform iterative looping by varying assigned weights until MSE value is less than 0.001 between predicted output and exact output. Besides that, the iterative process would also be ceased should the numbers of looping reaches 250,000 cycles. The final values of weights would be automatically saved within the ANN algorithm itself.

In the testing phase, the same set of data in training phase would be once again fed into the ANN for output prediction. MSE and linear correlation coefficient, *r* value would be calculated automatically to ensure the accuracy of the ANN system before further usage. The process of validation phase would be similar to testing phase. The only different was the data used in validation stage were never fed into the ANN system during both training and testing phase.

3.3 *Data Analysis*

From the IDARC-2D seismic analysis, seven important input parameters which were compulsory as the software input were taken as preliminary parameters for the ANN input. The parameters included age, height, and length, maximum number of bay as well as natural period of buildings. The other two parameters were seismic zone and ground acceleration. These seven parameters were considered as input for the ANN algorithm while the output would consist only of the damage index for each building. The preliminary overall diagram for the ANN algorithm is as illustrated in Figure 4.

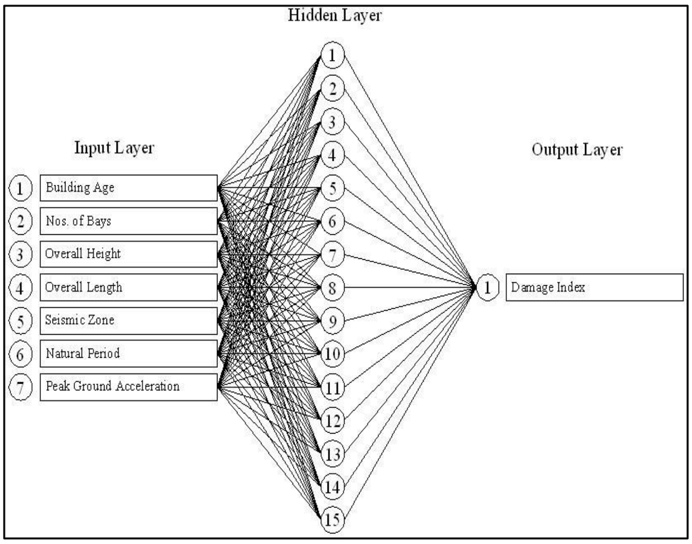

Figure 4. Preliminary diagram of ANN for prediction of building's damage index

There were altogether 164 sets of discrete input data for each of the seven identified parameters which would require to be classified into several different data class. The main purpose of this effort was to eliminate large amount of random data being fed as input into the ANN model. These parameters, together with their data classification details are, listed in Table 3.

3.4 *Training, Testing and Validation*

After the distribution of input and output parameters was determined, the data were divided into two groups; (a) training and testing stage and (b) validation phase. Among the 164 sets of data, 112 data were chosen randomly for the first stage while the remaining 52 sets of data were reserved for validation purposes. To ensure unbiased prediction by the neural network system, the validation data was not involved in training and testing phase. Details and information regarding the validation samples are contained in Table 4.

During the training process of the ANN, the influence of each input parameters towards the overall accuracy of the ANN model in prediction of damage index was studied by utilizing the mean square error analysis (MSE). The MSE for each model was calculated by varying the types of input parameters. Besides studying the optimum numbers of input parameters, the optimum numbers of hidden neuron was also studied using the same mathematical method by applying different quantity of hidden neurons.

Apart from using MSE analysis, the optimum numbers of input parameters as well as numbers of hidden neuron was also investigated during the testing and validation stages by utilizing the linear correlation coefficient, *r* method.

Table 4. Information of buildings for ANN accuracy validation

Building	*Length* (m)	*Height* (m)	*Bay* (nos.)	*Age* (years)	*Period* (s)
1	16.0	64.2	2	27	2.03
2	14.0	19.2	2	16	0.53
3	64.0	51.0	8	28	1.58
4	22.5	48.6	3	26	1.01
5	9.0	23.6	1	41	0.54
6	12.0	19.2	2	34	0.47
7	18.0	27.6	2	34	0.79
8	67.2	43.4	8	7	1.24
9	11.0	49.0	3	9	1.40
10	14.4	26.8	2	5	0.86
11	46.0	36.4	10	13	1.08
12	39.0	10.8	7	5	0.82
13	6.0	14.5	1	37	0.54

4 RESULTS

The results of MSE for obtaining the optimum numbers of input parameters and numbers of hidden neurons are shown in Table 5 and 6 below. Lower value of MSE reflects higher prediction accuracy of the ANN model.

Table 5. MSE values for different input parameters

Item	*Input Parameters*	*Mean Square Error(MSE)*
1	All input	0.027
2	Age excluded	0.041
3	Bay excluded	0.036
4	Height excluded	0.029
5	Length excluded	0.025
6	Zone excluded	0.031
7	Period excluded	0.037
8	G. acceleration excluded	0.032

Table 6. MSE values for different numbers of hidden neurons

Item	*Number of hidden neurons*	*Mean Square Error (MSE)*
1	1	0.057
2	5	0.088
3	10	0.027
4	15	0.027
5	20	0.095
6	25	0.115
7	30	0.372

From Table 5, it is clearly shown that the MSE value hit the lowest when the length of the building was eliminated by showing reading of 0.025. This value did not differ much with the MSE value of 0.027 when all input parameters were included. Meanwhile according to Table 6, the optimum numbers of hidden neurons was found to be 10 or 15 with the lowest MSE value of 0.027 for both cases.

On the other hand, the results of linear correlation coefficient, *r* method analysis to obtain the most optimum numbers of input parameters as well as numbers of hidden neurons are shown in Table 7 and 8. The closer the value of *r* is to 1, the more accurate it is for the ANN model.

Table 7. Linear correlation coefficient values for different input parameters

Item	*Input Parameters*	*Linear Correlation Coefficient, r*	
		Testing Phase	Validation Phase
1	All input	0.839	0.726
2	Age excluded	0.241	0.362
3	Bay excluded	0.520	0.714
4	Height excluded	0.286	0.461
5	Length excluded	0.342	0.274
6	Zone excluded	0.468	0.521
7	Period excluded	0.762	0.810
8	Ground acceleration excluded	0.802	0.621

Table 8. Linear correlation coefficient values for different numbers of hidden neurons

Item	*Number of Hidden Neurons*	*Linear Correlation Coefficient, r*	
		Testing Phase	Validation Phase
1	1	0.532	0.354
2	5	0.789	0.665
3	10	0.839	0.709
4	15	0.839	0.726

5	20	0.564	0.453
6	25	0.332	0.654
7	30	0.457	0.224

By comparing the MSE and *r* values of the four tables, the most optimum numbers of input parameters and numbers of hidden neurons are selected as follow; numbers of input parameter selected = 7; and numbers of hidden neurons selected = 15.

With this, the developed ANN model had achieved satisfactory percentage of accuracy in both testing phase and validation analysis, as illustrated in Figure 5 and 6.

5 CONCLUSIONS

An Artificial Neural Network (ANN) algorithm geared for buildings' damage index prediction due to seismic forces was successfully innovated. In the study, it was found out that the ANN gave highest accuracy when all seven preliminary identified input parameters were used together rather than eliminating any one of them. Meanwhile, the ANN system also achieved the highest accuracy by utilizing 15 numbers of hidden neurons.

The system managed to produce accurately 104 numbers of outputs among 112 building samples during testing phase, marking 93% of accurateness. However, the accuracy percentage in validation phase was only 75%, from which only 39 out of 52 building samples were being predicted accurately. From these percentages, it was clearly shown that the characteristic of input data being fed into the ANN system influenced the prediction results significantly. The distribution pattern and amount of these inputs in training phase played very important role in determining the overall accuracy of the ANN system.

This study focused on feasibility of adopting ANN in predicting structural damages due to seismic ground motions, as well as parametric investigation to determine the effect of different combinations of input parameters in affecting the prediction accuracy. Therefore, quantification of uncertainties which is often one of the key indicators used in conventional seismic risk assessment is not included in this study. The ANN directly predicted seismic damage of the frame structural system based on the input training data from the IDARC-2D analyses.

REFERENCES

Abdullah, C.S., Bakar, A.A., and Yusof, Y. (2000). Neural Network Modeling for In-Situ Concrete Strength. *Proceedings of the 14th Asia-Pacific Structural Engineering & Construction Conference 2000*, Malaysia, pp 223-236.

ATC-21. (1988). *Rapid Visual Screening of Buildings for Potential Seismic Hazards: a Handbook*. Washington DC: Applied Technology Council (FEMA).

ATC-22. (1989). *Rapid Visual Screening of Buildings for Potential Seismic Hazards: a Handbook.* Washington DC: Applied Technology Council (FEMA).

Ismail, R. (2008). *The Application of Artificial Neural Network in Seismic Evaluation of Buildings*. Master Thesis, Skudai: Universiti Teknologi Malaysia.

Kunnath, S.K., Reinhorn, A.M., and Lobo, R.F. (1992). "IDARC Version 3.0: A Program for the Inelastic Damage Analysis of Reinforced Concrete Structures." *National Center for Earthquake Engineering Research Report No. NCEER-92-0022*. New York: State University of New York at Buffalo.

Masters, T. (1993). *Practical Neural Network Recipes in C++*. San Diego CA: Academic Press.

Park, Y.J., Ang, A.H.S., and Wen, Y.K. (1984). "Seismic Damage Analysis and Damage Limiting Design of R/C Buildings." Civil Engineering Studies, Technical Report No. SRS 516, Urbana: University of Illinois.

Park, Y.J., Ang, A.H.S., and Wen, Y.K. (1987). "Damage-Limiting Aseismic Design of Buildings." *Earthquake Spectra*. Vol. 3, No.1, pp 1-26.

Tang Z. and Fishwick, P.A. (1993). "Feed Forward Neural Network as Models for Time Series Forecasting." *ORSA Journal of Computing 5 (A)*, pp 374-386.

Tsou, P. and Shen, M.H. (1993). "Structural Damage Detection and Identification Using Neural Network." *Proceedings of the 34th AIAA/ASME/ASCEAHS/ASC, Structural, Structural Dynamics and Materials Conference, AIAA/ASME Adaptive Structural Forum, Pt.5.*

Wu, X., Ghaboussi, J. and Garrett, J. (1992). "Use of Neural Network in Detection of Structural Damage." *Computer & Structure.* Vol. 42(4), pp 649-659.

Zhao, J., Ivan, J.N., and DeWold, J.T. (1998). "Structural Damage Detection Using Artificial Neural Networks." *Journal of Infrastructural Systems.* Vol. 13, No. 3, pp 182-189.

Table 3. Data classification for ANN

Parameter	Unit	Data Classification	Class Range	Frequency
INPUT				
X_1: Age	Year	Age Class 1	$X_1 \leq 10$	44
		Age Class 2	$10 < X_1 \leq 20$	32
		Age Class 3	$20 < X_1 \leq 30$	52
		Age Class 4	$30 < X_1 \leq 40$	20
		Age Class 5	$40 < X_1 \leq 50$	12
		Age Class 6	$50 < X_1 \leq 60$	4
X_2: No. of Bay	Number	1	1	24
		2	1	36
		3	1	56
		4	1	4
		5	1	0
		6	1	8
		7	1	4
		8	1	12
		9	1	4
		10	1	4
		11	1	8
		12	1	0
		13	1	0
		14	1	4
X_3: Height	Meter (m)	Height Class 1	$X_3 \leq 10$	12
		Height Class 2	$10 < X_3 \leq 20$	60
		Height Class 3	$20 < X_3 \leq 30$	28
		Height Class 4	$30 < X_3 \leq 40$	28
		Height Class 5	$40 < X_3 \leq 50$	16
		Height Class 6	$50 < X_3 \leq 60$	8
		Height Class 7	$60 < X_3 \leq 70$	8
		Height Class 8	$70 < X_3 \leq 80$	4
X_4: Length	Meter (m)	Length Class 1	$X_4 \leq 10$	44
		Length Class 2	$10 < X_4 \leq 20$	52
		Length Class 3	$20 < X_4 \leq 30$	24
		Length Class 4	$30 < X_4 \leq 40$	8
		Length Class 5	$40 < X_4 \leq 50$	16
		Length Class 6	$50 < X_4 \leq 60$	0
		Length Class 7	$60 < X_4 \leq 70$	8
		Length Class 8	$70 < X_4 \leq 80$	12
X_5: Seismic Zone	PGA (gal)	Seismic Zone 1	$30 < X_5 \leq 50$	40
		Seismic Zone 2	$50 < X_5 \leq 70$	72
		Seismic Zone 3	$70 < X_5 \leq 90$	8
		Seismic Zone 4	$90 < X_5 \leq 110$	4
		Seismic Zone 5	$110 < X_5 \leq 130$	0
		Seismic Zone 6	$130 < X_5 \leq 150$	5
X_6: Natural Period	Second (s)	Period Class 1	$X_6 \leq 0.5$	60
		Period Class 2	$0.5 < X_6 \leq 1.0$	60
		Period Class 3	$1.0 < X_6 \leq 1.5$	24
		Period Class 4	$1.5 < X_6 \leq 2.0$	14
		Period Class 5	$2.0 < X_6 \leq 2.5$	4

X_7: G. Acceleration	g	Acc. Class 1	0.05	41
		Acc. Class 2	0.10	41
		Acc. Class 3	0.15	41
		Acc. Class 4	0.20	41
OUTPUT				
Y_1: Damage Index	Damage Level	None	$Y_1 = 0$	23
		Slight	$0.00 < Y_1 \le 0.01$	1
		Minor	$0.01 < Y_1 \le 0.30$	124
		Moderate	$0.30 < Y_1 \le 1.00$	5
		Collapse	$Y_1 \ge 1.00$	11

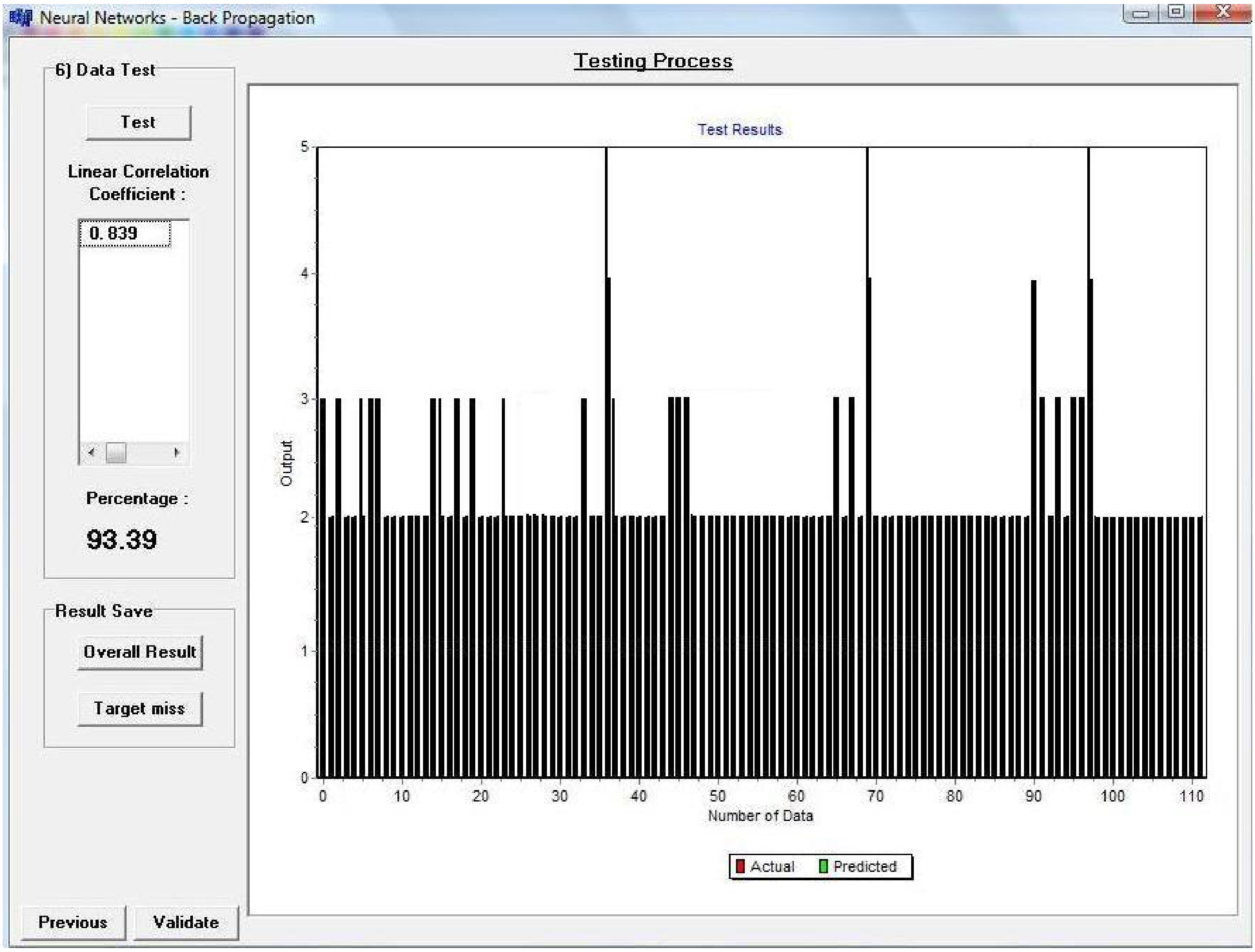

Figure 5. Screen shot of testing phase (93% accuracy percentage)

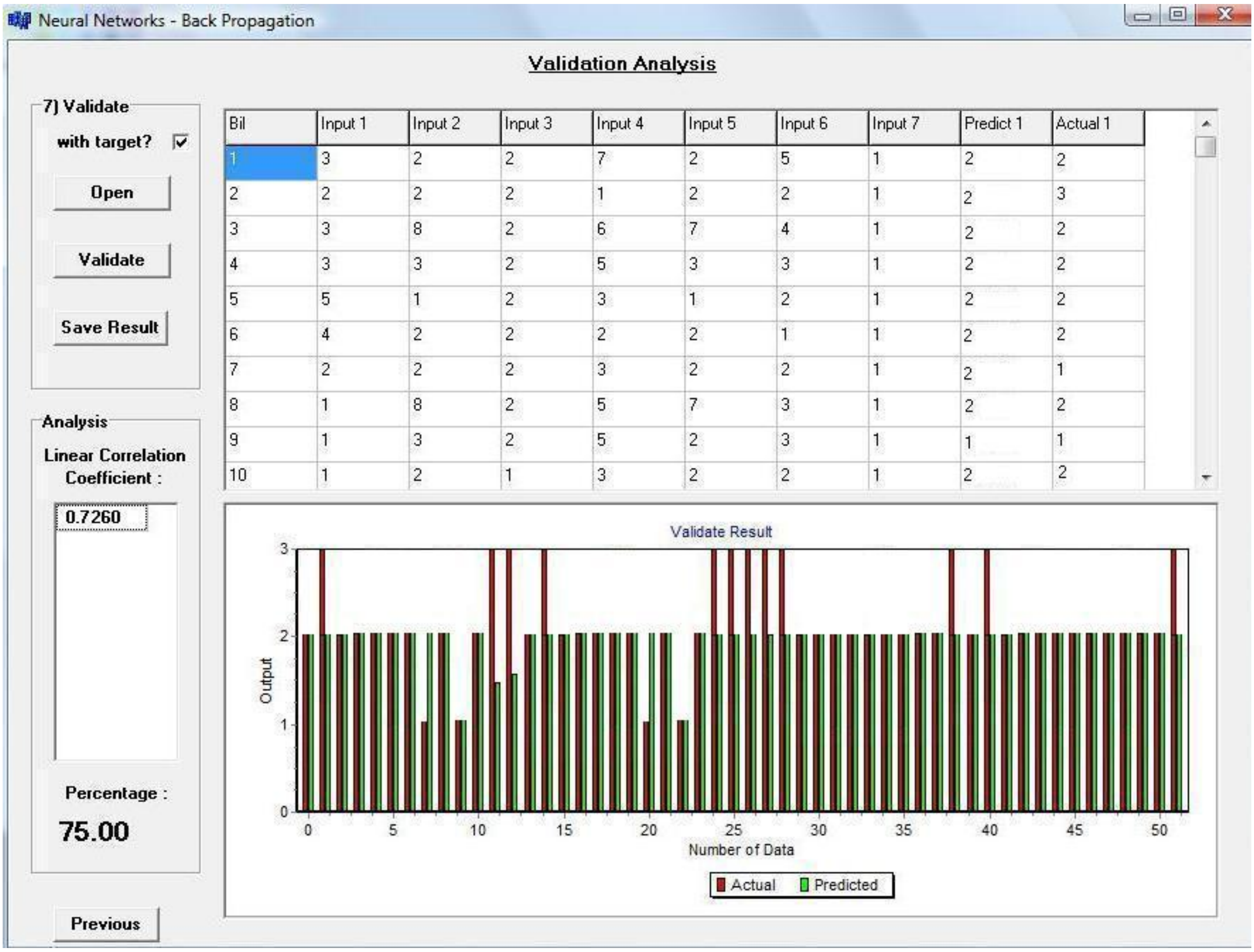

Figure 6. Screen shot of validation phase (75% accuracy percentage)