Proceedings of the World Congress on Engineering 2007 Vol I
WCE 2007, July 2 - 4, 2007, London, U.K.

Parallel Wavelet BP Neural Networks for Approximate Structural Dynamic Analysis

S. Gholizadeh, R. Kamyab, and C.K. Keong

Abstract—**dynamic analysis of structures for earthquake induced loads is very expensive in terms of the computational burden. In this study, to reduce the computational effort a new neural system which is called parallel wavelet back propagation (PWBP) neural networks has been introduced. Training of PWBP is implemented in two phases. In the first phase, the input space is classified by using competitive neural networks. In the second phase, one distinct WBP neural network is trained for each class. Comparison the numerical results obtained by PWBP with the corresponding ones obtained by the single WBP neural network reveals the better performance generality of PWBP.**

Index Terms—**back propagation, competitive, earthquake, neural network.**

I. Introduction

In the recent years, neural networks are considered as more appropriate techniques for solving the complex and time consuming problems. They are broadly utilized in civil and structural engineering applications. Determining the dynamic time history responses of structures for the earthquakes loadings is one of the time consuming problems with a huge computational burden. In the present study, neural networks are employed to predict the time history responses of structures. Some neural networks such as radial basis function (RBF), generalized regression (GR), counter propagation (CP), back propagation (BP) and wavelet back propagation (WBP) neural networks are used in civil and structural engineering applications [1-3]. As shown in Refs. [2-3] the performance generality of WBP for approximating the structural time history responses is better than that of the RBF, GR, CP and BP neural networks. Therefore, in this study we have focused on WBP neural networks and its improvements. The most important phase in the neural networks training is data generation. As emphasized in the relevant professional literatures such as Ref. [4] there is no explicit method to select the training samples and therefore this job is usually accomplished on the random basis. Therefore, in the large scale problems selection of proper training data may require significant computer effort. Also, in the case of such problems, to train a robust neural network, many training samples must be selected. In the present paper, we introduce a new neural system for eliminating the main difficulties occurred in training mode of WBP neural networks. The new system is designed in two main phases. In the first phase, the input space is classified based on one criterion using a competitive neural network. In the second phase, one distinct WBP neural network is trained for each class using data located. In this manner, a set of parallel WBP neural networks are substituted with a single WBP neural network. The neural system is called parallel wavelet back propagation (PWBP) neural networks. The numerical results indicate that the performance generality of PWBP is better than that of the single WBP neural network.

Manuscript received February 26, 2007. This work was supported by the Iranian Academic Center Education, Culture and Research.

S. Gholizadeh is with the Civil Engineering Department, Iranian Academic Center for Education, Culture and Research, Kerman, IRAN (corresponding author to provide phone & fax: +98-341-3221443; +98-341-3221442; +98-341-3221441; e-mail: gholizadeh@acecr.ac.ir).

R. Kamyab., is with the Civil Engineering Department, Iranian Academic Center for Education, Culture and Research, Kerman, IRAN. (e-mail: r_kamyab@acecr.ac.ir).

C.K. Keong is with the School of Civil Engineering, University of Science Malaysia, Engineering Campus Transkrian, Pulau Pring, Malaysia (e-mail: ckchen@kck.usm.my).

II. Back propagation neural networks

Back Propagation was created by generalizing the Widrow-Hoff learning rule to multiple layer neural networks and nonlinear differentiable transfer functions. Input vectors and the corresponding target vectors are used to train a network until it can approximate a function, associate input vectors with specific output vectors. Neural networks with a sigmoid layer and a linear output layer are capable of approximating any function with a finite number of discontinuities. Standard back propagation is a gradient descent algorithm, as is the Widrow-Hoff learning rule, in which the neural network weights are moved along the negative of the gradient of the performance function. The term back propagation refers to the manner in which the gradient is computed for nonlinear multilayer neural networks. There are a number of variations on the basic algorithm that are based on other standard optimization techniques, such as conjugate gradient and Newton methods. In this study we have employed scaled conjugate gradient (SCG) algorithm that was developed by Moller [5].

III. Fundamentals of wavelet theory

Wavelet theory is the outcome of multi-disciplinary endeavours that brought together mathematicians, physicists and engineers. This relationship creates a flow of ideas that goes well beyond the construction of new bases or transforms. The term of wavelet means a little wave. A function $h \in L^2(R)$ (the set of all square integrable or a finite energy function) is called a wavelet if it has zero average on $(-\infty, +\infty)$ [6]:

$$\int_{-\infty}^{+\infty} h(t)dt = 0 \tag{1}$$

This little wave must have at least a minimum oscillation and a fast decay to zero in both the positive and negative directions of its amplitude. These three properties are the Grossmann-Morlet admissibility conditions of a function that is required for the wavelet transform. The wavelet transform is an operation which transforms a function by integrating it with modified versions of some kernel functions. The kernel function is called the mother wavelet and the modified version is its daughter wavelet. A function $h \in L^2(R)$ is admissible if:

$$c_h = \int_{-\infty}^{+\infty} \frac{|H(\omega)|^2}{|\omega|} d\omega = 0 \tag{2}$$

where $H(\omega)$ is the Fourier transform of h(t). The constant c_h is the admissibility constant of the function h(t). For a given h(t), the condition $c_h < \infty$ holds only if $H(0) = 0$.

The wavelet transform of a function $h \in L^2(R)$ with respect to a given admissible mother wavelet h(t) is defined as:

$$W_f(a,b) = \int_{-\infty}^{+\infty} f(t) h^*_{a,b}(t) dt \tag{3}$$

where * denotes the complex conjugate. However, most wavelets are real valued.

Sets of wavelets are employed for approximation of a signal and the goal is to find a set of daughter wavelets constructed by dilated and translated original wavelets or mother wavelets that best represent the signal. The daughter wavelets are generated from a single mother wavelet $h(t)$ by dilation and translation [7]:

$$h_{a,b}(t) = \frac{1}{\sqrt{a}} h(\frac{t-b}{a}) \tag{4}$$

Where $a > 0$ is the dilation factor, b is the translation factor.

IV. Wavelet Neural Networks

Wavelet neural networks (WNN) employing wavelets as the activation functions recently have been researched as an alternative approach to the neural networks with sigmoidal activation functions. The combination of wavelet theory and neural networks has lead to the development of WNNs. WNNs are feed forward neural networks using wavelets as activation function. In WNNs, both the position and the dilation of the wavelets are optimized besides the weights. Wavenet is another term to describe WNN. Originally, wavenets did refer to neural networks using wavelets. In wavenets, the position and dilation of the wavelets are fixed and the weights are optimized [6].

V. Wavelet Back Propagation Neural Networks

BP network is now the most popular mapping neural network. But it has few problems such as trapping into local minima and slow convergence. Wavelets are powerful signal analysis tools. They can approximately realize the time-frequency analysis using a mother wavelet. The mother wavelet has a square window in the time-frequency space. The size of the window can be freely variable by two parameters. Thus, wavelets can identify the localization of unknown signals at any level. Activation function of hidden layer neurons in BP neural network is a sigmoidal function shown in Fig.1a. To design wavelet back propagation (WBP) neural network we substitute hidden layer sigmoidal activation function of BP with POLYWOG1 wavelet [7]:

$$h_{POLYWOG\ 1}(t) = \sqrt{\exp(1)}(t)\exp(-t^2/2) \tag{5}$$

Plot of POLYWOG1 with a=1 and b=0, is shown in Fig.1.b. In the resulted WBP neural network, the position and dilation of the wavelets as activation function of hidden layer neurons are fixed and the weights of network are optimized using the SCG algorithm. In this study, we obtain good results considering b = 0 and a = 2.5. The activation function of the hidden layer neurons is as (6).

$$h_{POLYWOG\ 1}(t) = \sqrt{\exp(1)}(t/2.5)\exp(-(t/2.5)^2/2) \tag{6}$$

Therefore, WBP is a modified back propagation neural network with POLYWOG1 hidden layer neurons activation function. And adjusting the weights of the neural network is performed using SCG algorithm. Typical topology of WBP is shown in Fig.2.

VI. Competitive Neural Networks

Some applications need to group data that may, or may not be, clearly definable. Competitive neural networks can learn to

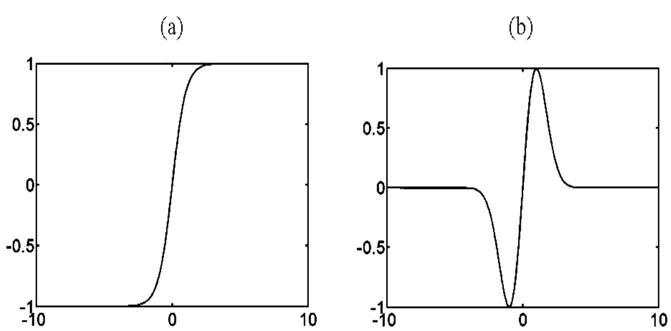

Fig.1:a) Sigmoidal function, b) POLYWOG1 wavelet

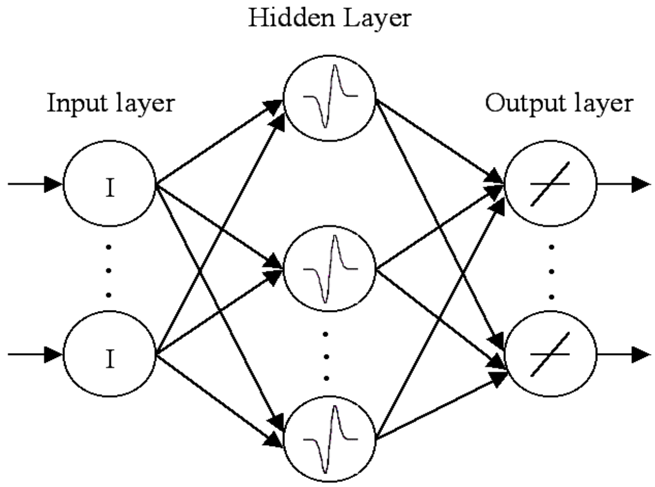

Fig.2: Typical topology of WBP

detect regularities and correlations in their input and adapt their future responses to that input accordingly. The neurons of competitive networks learn to recognize groups of similar input vectors. A competitive network automatically learns to classify input vectors. However, the obtained classes by the competitive network depend only on the distance between input vectors. If two input vectors are very similar, the competitive network probably will put them in the same class. There is no mechanism in a strictly competitive network design to say whether or not any two input vectors are in the same class or different classes. A competitive network simply tries to identify groups as best as they can. Training of competitive network is based on Kohonen [10] self-organization algorithm. A key difference between this network and many other networks is that the competitive network learns without supervision. During training the weights of the winning neuron are updated according to:

$$w_{ij}(k+1) = w_{ij}(k) + \alpha[x_j(k) - w_{ij}(k)] \quad (7)$$

where w_{ij} is the weight of competitive layer from input i to neuron j, x_j is jth component of the input vector, α is learning rate and k is discrete time.

VII. Parallel Wavelet Back Propagation Neural Networks

As mentioned previously, in the case of large problems to train a neural network with acceptable performance generality, it is quite necessary that an adequate number of training data to be selected. Therefore, too much computer effort is required in the training phase. To attain appropriate generalization spending low effort we propose the PWBP neural networks.

At first, the selected input-target training pairs are classified in some classes based on a specific criterion. In other words, the input and target spaces are divided into some subspaces as the data located in each subspace have similar properties. Now we can train a small WBP neural network for each subspace using its assigned training data. Considering the mentioned simple strategy a single WBP neural network which is trained for all over the input space is substituted with a set of parallel WBP neural networks as each of them is trained for one segment of the classified input space.

In PWBP, each WBP neural network has specific dilation factor which may differ from that of the other WBP neural networks. Therefore performance generality of PWBP neural networks is higher than that of the single WBP neural network.

Improving generalization process of PWBP is performed very rationally and economically in comparison with that of the single WBP neural network. In other words, improving generalization process and retraining of some small parallel WBP neural networks have low effort with respect to those of the single WBP neural network. Furthermore, it is very probable that some of the parallel WBP neural networks of PWBP require no improving generalization..

Selection a proper criterion for classification of the input space depends on the nature of the problem and its variables thus recognition of the effective arguments to select an efficient criterion has very significant influence on the generality of PWBP. Determination the number of the classes depends on the complicacy and size of the input space and there are no special criteria for this mean.

VIII. Error Estimation

In the present study to evaluate the error between exact and approximate results, the root mean squared error (RMSE) is calculated.

$$\text{rmse} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \tilde{x}_i)^2} \quad (8)$$

where, x_i and $\tilde{x}_i$ are the ith component of the exact and approximated vectors, respectively. n is the vectors dimension.

To measure how successful fitting is achieved between exact and approximate responses, the Rsquare statistic measurement is employed. A value closer to 1 indicates a better fit.

$$\text{Rsquare} = 1 - \frac{\sum_{i=1}^{n}(x_i - \tilde{x}_i)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad (9)$$

where, $\bar{x}$ is the mean of exact vectors component.

IX. Numerical Results

A. *Example 1*

A three stories steel frame structure is shown in Fig.3. Rigid diaphragms are assigned to the roof of each story. Cross sections of columns and beams are selected from the wide flange sections available in European profile list. Spans in x and y directions are 4 m. Height of each storey is 3 m. Distributed load of 700 kg/m^2 are considered on the

ISBN:978-988-98671-5-7

diaphragms. In order to practical demands and simplify the time history analysis, cross-sections of the columns are selected from the profiles listed in Table 1. Also, a profile of IPE 300 is assigned to the all beams.

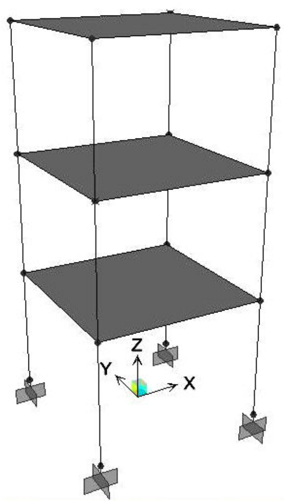

Fig. 3: Steel Frame

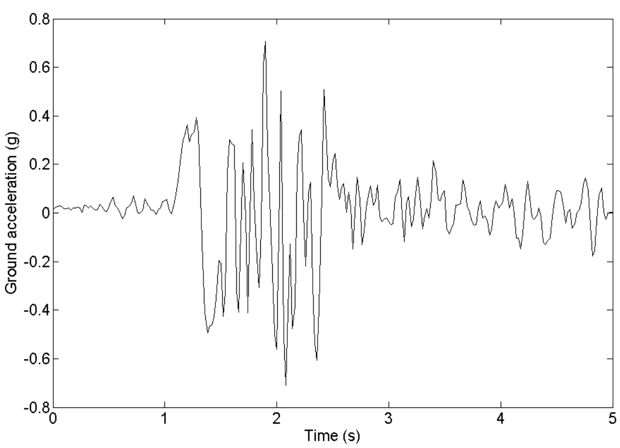

Fig. 4: The Naghan earthquake records (Iran 1977)

Due to practical demands unique cross-sectional properties is considered for all columns in each story.

Table 1: Available properties

No	Columns
1	HE 160-M
2	HE 200-M
3	HE 240-M
4	HE 280-M
5	HE 320-M
6	HE 360-M
7	HE 450-M

B. Data generation

In this problem input space consists of three natural periods of the structures and corresponding time history responses of the third story in x direction with respect to the Naghan earthquake are considered as target space. A total number of 46 structures are randomly generated, based on cross-sectional properties, and are analyzed for the earthquake induced loads. From which 31 samples are used for training and 15 samples are employed to test the performance generality of the neural networks. In this study to design the neural networks MATLAB [8] is employed and dynamic analysis of the structures is performed using SAP2000 [9].

C. Training the single WBP neural network

A single WBP neural network is trained using the mentioned training set. In this network 14 WBP neurons and 250 linear neurons are assigned to the hidden and output layers, respectively. As shown in Table 2 the performance generality of the neural network is not very good. To attain appropriate performance generality PWBP is employed.

D. Training PWBP neural networks

As mentioned in section V., we need a criterion to classify the input space. Here, we employ natural periods of the structure as the classification criterion. It is obvious that the structures with similar natural periods yield the same patterns for dynamic structural responses. To classify the input space a competitive neural network is trained. In this problem we obtain the best results by choosing three classes. After classification, a distinct WBP neural network can be trained for each class. The results of testing the WBP and PWBP neural networks are summarized in Table 2.

Table 2: Rsquare and rmse of the Single WBP and PWBP neural networks

Test	Single WBP		PWBP		
	Rsquare	rmse	Rsquare	rmse	Class
1	0.990	0.197	0.982	0.273	2
2	0.726	0.976	0.988	0.210	2
3	0.957	0.366	0.967	0.320	2
4	0.937	0.398	0.985	0.198	2
5	0.934	0.438	0.985	0.215	2
6	0.854	0.711	0.997	0.096	2
7	0.956	0.458	0.992	0.195	3
8	0.993	0.176	0.993	0.180	1
9	0.934	0.535	0.995	0.103	1
10	0.994	0.118	0.999	0.059	1
11	0.949	0.490	0.984	0.280	1
12	0.807	0.724	0.970	0.365	1
13	0.992	0.194	0.988	0.241	3
14	0.981	0.297	0.992	0.198	3
15	0.982	0.279	0.999	0.067	3
Average	0.932	0.424	0.988	0.200	-

The displayed results in the above table indicate that the performance generality of PWBP neural networks is better than that of the single WBP network. Rsquare and rmse of the parallel WBP neural networks are shown in Tables 3 and 4, respectively. The parallel WBP neural networks of class 1, 2, and 3 are trained with 3, 3, and 2 hidden layer neurons, respectively.

To improve the performance generality of the single WBP neural network, the number of training data should be frequently increased until the acceptable accuracy is obtained.

Table 3: Rsquare of PWBP neural networks

Metric	Class		
	1	2	3
Rsquare	0.993	0.945	0.992
	0.995	0.991	0.988
	0.999	0.972	0.992
	0.984	0.985	0.999
	0.970	0.985	-
	-	0.997	-
Average	0.988	0.980	0.993

Table 4: Rsquare of PWBP neural networks

Metric	Class		
	1	2	3
rmse	0.180	0.273	0.194
	0.103	0.210	0.241
	0.059	0.320	0.198
	0.280	0.198	0.067
	0.365	0.215	-
	-	0.096	-
Average	0.197	0.219	0.175

E. Example 2

A ten bar steel truss is shown in Fig. 5. Cross-sectional areas of the members are selected from the pipe sections available in European profile list. The truss was subjected to the El Centro earthquake records (S-E 1940), displayed in Fig. 6, which is effectively used in x direction. Span in x direction and height of the truss is 3 m and 6 m, respectively. The mass of 5000 kg is lumped at each free node.

In order to simplify the time history analysis, 8 types of cross-sectional areas are considered for the truss elements which are displayed in Table 5.

Because of zero internal stresses of elements 5 and 6 under the earthquake loading, a minimum cross sectional area of 0.515 cm^2 was assigned to them. Due to simplicity and practical demands, the truss members are divided into 6 groups based on cross-sectional areas, shown in Table 6.

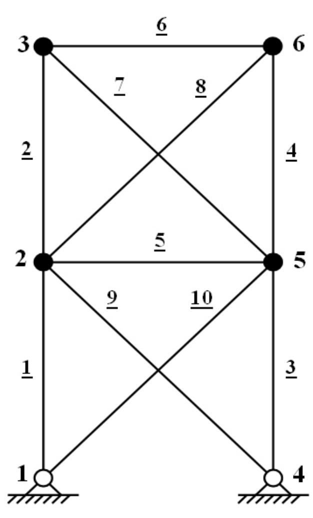

Fig. 5: 10-bar steel truss

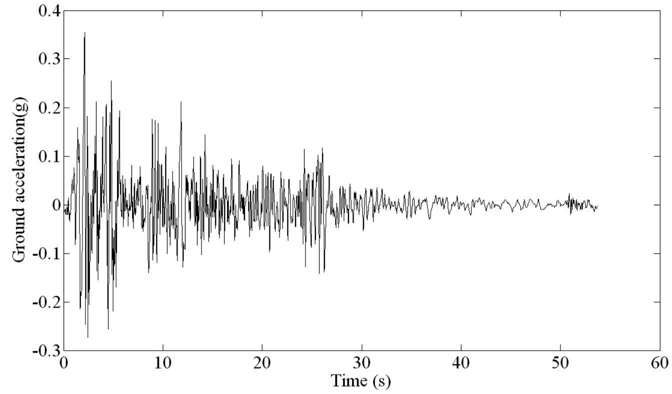

Fig. 6: The El Centro (S-E 1940) records

Table 5: Available pipe profiles

No.	Area (cm^2)
1	12.5
2	13.7
3	17.2
4	25.1
5	27.2
6	31.1
7	50.0
8	52.7

Table 6: Element groups of the 10-bar truss

Group	1	2	3	4	5	6
Members	1	2	3	4	7 8	9 10

F. Data generation

In this problem the input space consists of four higher natural periods of the structures and corresponding time history responses of node 6 in x direction against the El Centro earthquake are considered as target space. At first a total number of 240 structures are randomly generated based on cross-sectional areas, from which 200 samples are used for training and 40 ones are employed to test the neural networks performance generality.

G. Training the single WBP neural network

To train the single WBP neural network 40 WBP neurons and 2688 linear neurons are assigned to hidden and output layers, respectively.

H. Training PWBP neural network

In this problem we obtain the best results with five classes. The number of training data located in class 1 to 5 is 30, 41, 37, 40 and 52, respectively. In this example the parallel WBP neural networks have 4, 5, 4, 5 and 5 WBP neurons. The test results of the single WBP and PWBP neural networks are summarized in Table 7. The results indicate that the performance generality of PWBP is better than that of the single WBP. Rsquare and rmse of the parallel WBP neural networks are shown in Tables 8 and 9, respectively.

Table 7: Table 2: Rsquare and rmse of the Single WBP and PWBP neural networks

Test	Single WBP		PWBP		
	Rsquare	rmse	Rsquare	rmse	Class
1	0.872	0.220	0.988	0.066	1
2	0.839	0.200	0.994	0.036	1
3	0.646	0.255	0.989	0.043	1
4	0.918	0.164	0.992	0.049	1
5	0.872	0.190	0.995	0.036	1
6	0.946	0.106	0.992	0.039	1
7	0.643	0.685	0.873	0.158	1
8	0.753	0.225	0.990	0.045	1
9	0.924	0.271	0.995	0.067	2
10	0.959	0.178	0.995	0.060	2
11	0.899	0.331	0.992	0.088	2
12	0.918	0.250	0.964	0.165	2
13	0.933	0.256	0.995	0.064	2
14	0.709	0.460	0.925	0.233	2
15	0.941	0.247	0.996	0.059	2
16	0.628	0.544	0.767	0.429	2
17	0.969	0.235	0.995	0.090	3
18	0.919	0.357	0.901	0.395	3
19	0.394	1.126	0.978	0.214	3
20	0.983	0.144	0.991	0.102	3
21	0.986	0.160	0.998	0.054	3
22	0.939	0.300	0.998	0.047	3
23	0.982	0.179	0.996	0.085	3
24	0.946	0.304	0.996	0.082	3
25	0.946	0.278	0.995	0.084	4
26	0.970	0.231	0.997	0.060	4
27	0.967	0.213	0.998	0.041	4
28	0.934	0.339	0.996	0.075	4
29	0.960	0.238	0.997	0.054	4
30	0.946	0.268	0.999	0.030	4
31	0.960	0.235	0.993	0.098	4
32	0.957	0.220	0.995	0.071	4
33	0.698	0.519	0.966	0.174	5
34	0.959	0.139	0.989	0.072	5
35	0.764	0.489	0.930	0.268	5
36	0.882	0.249	0.975	0.113	5
37	0.742	0.462	0.985	0.108	5
38	0.933	0.251	0.989	0.101	5
39	0.943	0.201	0.967	0.151	5
40	0.902	0.309	0.992	0.087	5
Average	0.875	0.300	0.976	0.108	-

Table 8: Rsquare of PWBP neural networks

Metric	Class				
	1	2	3	4	5
Rsquare	0.0659	0.0677	0.0899	0.0841	0.1737
	0.0361	0.0596	0.3954	0.0605	0.0719
	0.0432	0.0880	0.2137	0.0412	0.2638
	0.0492	0.1654	0.1023	0.0748	0.1133
	0.0364	0.0640	0.0542	0.0543	0.1082
	0.0394	0.2334	0.0467	0.0304	0.1006
	0.1581	0.0588	0.0849	0.0975	0.1511
	0.0449	0.4290	0.0819	0.0711	0.0872
Average	0.0592	0.1457	0.1336	0.0642	0.1337

Table 9: rmse of PWBP neural networks

Metric	Class				
	1	2	3	4	5
rmse	0.9886	0.9953	0.9954	0.9950	0.9662
	0.9947	0.9950	0.9013	0.9979	0.9891
	0.9898	0.9928	0.9782	0.9988	0.9307
	0.9926	0.9641	0.9915	0.9967	0.9756
	0.9953	0.9958	0.9984	0.9979	0.9859
	0.9925	0.9253	0.9985	0.9993	0.9892
	0.8732	0.9966	0.9960	0.9932	0.9678
	0.9901	0.7679	0.9961	0.9956	0.9922
Average	0.9771	0.9541	0.9819	0.9968	0.9746

X. Conclusion

An efficient method is introduced to predicting the time history responses of structures combining competitive and wavelet back propagation neural networks. The resulted neural system is called parallel wavelet back propagation (PWBP) neural networks. Two numerical examples are considered to demonstrate the computational advantages of the PWBP with respect to the single WBP neural networks. The numerical results reveal that PWBP neural networks can be effectively employed to predict the time history responses of structures for earthquake loading.

References

[1] E. Salajegheh, S. Gholizadeh, "Optimum Design of Structures by an Improved Genetic Algorithm Using Neural Networks", *Advances in Engineering Software*, Vol.36, No.11-12, 2005, pp. 757-767.

[2] E. Salajegheh, A. Heidari, "Optimum design of structures against earthquake by wavelet neural network and filter banks", *Earthquake Engineering and Structural Dynamics*, Vol.34, No.1, 2005, pp. 67-82.

[3] E. Salajegheh, S. Gholizadeh, "Comparison of RBF, GR and BP with wavelet back propagation networks in approximation dynamic analysis of structures", 7th International Congress on Civil Engineering, 2006, Tehran, Iran.

[4] M.Y. Rafiq, G. Bugmann, D.J. Easterbrook, "Neural network design for engineering applications", *Computers & Structures*, Vol.79, No.17, 2001, pp. 1541-1552.

[5] M.F. Moller, "A scaled conjugate gradient algorithm for fast supervised learning", *Neural Networks,* Vol.6, No.4, 1993, pp. 525-533.

[6] M. Thuillard, *A Review of Wavelet Networks, Wavenets, Fuzzy Wavenets and their Applications*, ESIT, Aachen, Germany, 2000.

[7] G. Lekutai, *Adaptive Self-Tuning Neuro Wavelet Network Controllers*, Ph.D thesis, Virginia Polytechnic Institute and State University, The Electrical Engineering Department, 1997.

[8] The Language of Technical Computing, "*MATLAB*", Math Works Inc. 2004.

[9] SAP2000 integrated finite element analysis and design of structures. Computers and Structures, Inc., Univ. Ave., Berkeley, CA., USA, 2001.