Application of Genetic Algorithm in Improvement of Efficiency of Neural Network

In Forecasting of Consumed Load Peak in West of Iran

Mahdi Goodarzi, Mahmud Goudarzi
Electrical engineering Department, Toyserkan Branch, Islamic Azad University, Iran
Info.goodarzi@gmail.com; goudarzian@gmail.com

***Abstract*-In this article application of genetic algorithm is discussed in order to improve the response of MLP neural network method for forecasting consumed load peak. In this method genetic algorithm is used which has more power in finding total minimum in comparison with algorithms which are based on Newton. This method could find consumed load peak of west of Iran more precise than other models with MRE% (Mean relative error percent) equal 1.207. Several models of ANN with a hidden layer have been checked and the best of them was chosen as the best structure with 3 neutrons. Data scattering test was done in order to correct choosing of test, train and validation data and this matter confirmed correctness of choosing data collection. For guaranty the authority of outcomes of the model, ±40MW considered after checking graph as error span, in a way that real data will be followed.**

Keywords-genetic algorithm; neural network; neural-genetic; consumed load peak

I. INTRODUCTION

Electrical load forecasting has been considered by the industrial researchers and university scholars since many years ago because of its important role in effective and economical operations.

Error in forecasting may result in too much risk or too much conservative timing and this matter can result in unwanted economical penalties. For example, higher forecasting may lead to production of more amount of electrical power and as a result, generated energy will be more than real demand, and lower forecasting may result in inability in production of energy and in both cases higher operational cost will be imposed [1]. With regards to the necessarily of amendment of consumption model in IRAN, there is need to the methods for forecasting of different consumers' future consumption. For forecasting we can use classic methods such as statistical techniques like regression, time series and calculation intelligence like neural networks or fuzzy logic. Since 1990 which artificial neural networks with abbreviation of ANN or NN were used in the forecasting of load, we can dare to say that this method was the most used one in the field of load forecasting [2]. Totally neural networks are formed form neurons and connections between them. Weights and Biases which are placed in neurons are optimized in order to teach the network for accessing minimum error. Normal methods in this section use error minimizing algorithms based on Newton methods. Weakness of these methods is not finding total optimum and this causes to decrease efficiency of the network. At the time being, genetic algorithm method is a very suitable method for finding original optimum. This algorithm is based on Darwin Evolution Theory based upon compatibility with environment. So many researches have been done with this method in order to remove weaknesses and deficiencies of intelligent techniques [3-12].

In this article combination of artificial neural network and genetic algorithm is used in order to predict peak of consumed load in west of Iran. Usual systems need weather information in order to predict load and lack of a reliable system for predicting weather situation results in too much complexity of the system [13, 14]. Published articles in this field [15] show that out of 22 articles in this filed 13 articles have considered temperature and in 3 cases they have used more information about weather and in three other cases information about loading has been used. But in the cases in which only loading information has been used as input, forecasting was along with too much error in such a way that it cannot be considered as a reliable system for load forecasting.

In this paper only the previous values of load peak have been used as input for forecasting load peak values in the future and it has been showed that the accuracy of this method in forecasting of load peak is at an acceptable level. With regards to above mentioned cases, values of load peak within previous week and values of loading peak on the same day in previous two years and an index which determines kind of the day are given to the system as input.

II. NETWORK STRUCTURE

A. *Multi-layer Perceptron Network*

Fig 1 shows one layer of multi-layer perceptron (MLP)

with R units input ($P_1,\dots,P_R$) and S neurons. In this network each input member from P vector is connected to each input neuron through weight matrix of W. Ith neuron has an accumulating unit which accumulates weighted inputs and Biases as an output of scalar n(i). Different n(i)s form a pure input vector n. Eventually outputs of neuron layer form column vector a, output of equation shown in equation (1).

$$n_j = \sum_{i=1}^{R}(p_i w_{ij} + b_j) \quad , j = 1,2,\dots S \qquad (1)$$

DOI 10.5963/JCET0201008

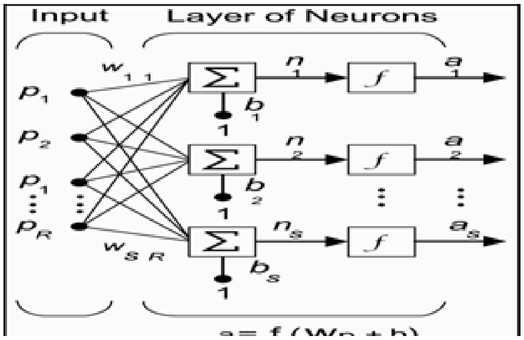

Fig. 1 Mono-layer network with R input and S neuron in intermediate layer

W_{ij} is connection weights, P_i is input unit and n_j is output unit and b_j is the Bias related to jth cell. The role of Bias is increasing or decreasing weighted collection and it helps the network to know present models better. Final output of the network (a) is calculated through equation (2) by using activation function of f(x).

$$a_S = f(n_S) \qquad (2)$$

in which calculated weights from training are saved and it prepares the information for future use.

B. Structure of Input Information

The first stage in designing an intelligent system is providing an efficient combination of inputs which have linear independence from each other and secondly each of them have useful information about the past time of the system. Linear independence of the data from each other means that if we assume X as the matrix of the input data, the condition of $|X*X'|\neq 0$ should be met. After reaching a 10*10 matrix from input data above condition was also met.

With regards to fig 2 we observe that change trends of consumed load peak in the west province of IRAN are almost weekly. It means if we draw graph of changes of consumed load peak based on week days, consumption growth starts since the beginning of the week and it reaches to its minimum value at the end of the week. However relative falling occurs on national holidays but the least consumption value is at the ending days of the week.

At the beginning and end of each week we observe different minimum and maximum that ascending or descending trend of these have different nature in comparison with other weeks. Therefore we have used values of load peak of 7 previous days and peak value on the same day in 2 previous years and another parameter which shows type of day in order to reach a better understanding of the previous influence of load.

The training data is obtained from the west side of Iran Electricity Distribution Corporation at 2009-2010.

With regards to different nature of week days and

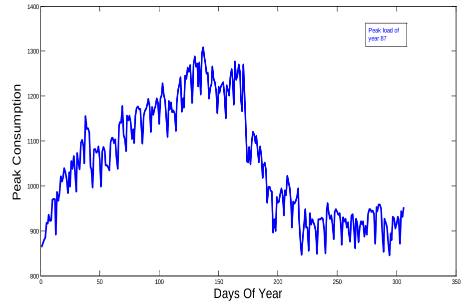

Fig. 2 Variation of peak load consumption in 2009-2010

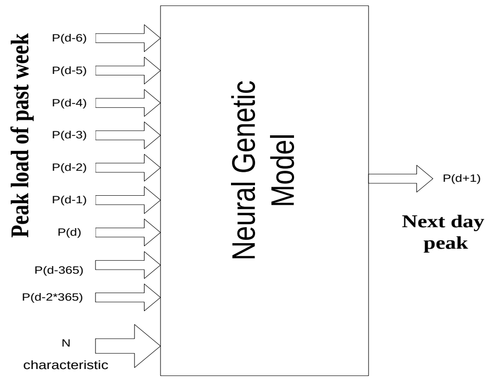

Fig. 3 Simplified scheme of the inputs and outputs and used neural network.

consequently their incongruent influence on consumption of electrical load, we have divided days into four types such as normal working days like Sunday to Wednesday, high consumption days like Saturday, part time days like Thursday and eventually national holidays and Fridays and we have respectively allocated 1,2,3, and 4 to them. Moreover one of the issues which is very useful in the precision of forecasting consumed load peak is scaling the data. Here input data have been converted into the scale between (1, 0) which makes the education of the network much easier. Equation (3) has been used in order to implement the scaling.

$$P_n = (P-P_{mean}) / P_{std} \qquad (3)$$

Where P_{mean} is mean of each row of p and P_{std} is standard deviation of each row of P.

Standard deviation (Std) of each parameter of X is obtained through equation (4).

$$std = \left(\frac{1}{n} \sum_{i=1}^{n} (x_i - \ddot{x})^2 \right)^{1/2} \qquad (4)$$

In which

$$\ddot{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \qquad (5)$$

where n shows number of samples.

DOI 10.5963/JCET0201008

C. Genetics Algorithm

Problems in determining the parameters of ANN model include choosing input elements, primary weights matrix and educating parameters (and the number of repetitions, etc).

Genetic algorithms (GAs) are a common method in solving optimizing issues. GAs was first suggested by John Holland et al. in 1992 which was based upon the genetic system of living beings according to their biological behavior. The idea of genetic algorithm is suggested based on the genetic and developmental systems in living beings and its purpose is to achieve an optimized result for a given problem; in a way that any natural organism has its own genetic information, those who can accommodate with their environment can transfer their genetic information to the upcoming generations through reproduction. Genetic mutation through genetic operators and mutation and Crossover of two chromosomes generates variety. In GA, finding the result to solve a problem is conducted through comparing the obtained result with experimental data. The obtained result is called Fitness whose name is drawn from fitting an organism in its environment.

The first step in solving the problem through genetic algorithm is coding input data in a special basis (usually 0 & 1 binary), these strings are called chromosomes. In solving the problems through genetic algorithm, a lot of chromosomes are produced randomly and after the application of genetic operators on them, their fitness is compared to the favorite result and finally after several phases a result (chromosome) is produced with high degree of fitness. Several methods are suggested for turning a series of chromosomes into a suitable result and here we apply a simple method [16].

D. Optimizing Weights and Biases through Using GA

Genetic operation using colony search technology can be avoided getting into local extreme point. This method can be helpful in reaching network optimized weights and biases. The whole operation is divided into 6 sections as follows.

1) Coding of Weights and Biases

Input data must be turned into the data that can be used in genetic algorithm (chromosomes and genes) and all of them must be coded in to the binary. The vector determining chromosomes is stated by equation 6 in general.

$$W=[W_1, W_2, \dots, W_n, \theta_1, \theta_2, \dots, \theta_n] \quad (6)$$

In which w_i is the related gene to ith weight and θj is the related gene to jth bias within an assumed chromosome.

2) The Primary Population

After coding operation on the weights and biases of each chromosome in the network, algorithm randomly makes a primary population. Algorithm begins a repeated search by using the primary population as the starting point. Finally, the population size, selection, crosses over and mutation values are determined in the order of 500, 0.8, 0.6, 0.001 through experiments [16, 17].

3) Fitness Function

Fitness function is an important basis in evaluating the members of the population. The most common Fitness function in training developmental network is the mean of the square error between prediction data and goal data. For optimizing generalization, the Fitness function could be corrected. This is done by adding a norm to the Fitness function that equals the mean aggregate of the square of network weights and biases. This function is illustrated in equation 7.

$$msereg = \gamma mse + (1-\gamma) msw \quad (7)$$

In which γ is efficiency coefficient. The amounts of mse and msw are calculated according to equation 8.

$$mse = \frac{1}{n}\sum_{i=1}^{n}(t_i - a_i)^2 \quad (8)$$

$$msw = \frac{1}{n}\sum_{j=1}^{n} w_j^{\,2} \quad (9)$$

In which ti is the actual data, ai is the predicted data, w is the weights matrix and n is the number of samples.

It needs to be mentioned that the amount of the γ parameter is calculated 0.5 through experiments. Using this parameter results in having a network with more reliable weights and biases and leads to a network with smoother results and consequently to a network with less over fit probability.

4) Selection Operation

This is done for selecting people with more fitness among the population. This operation provides the opportunity for the production of the next generation. In the present research, the Roulette wheel selection method is chosen for selecting new people. The probability of selecting people is as equation 10.

$$P_i = \frac{F_i}{\sum_{i=1}^{n} F_i} \quad (10)$$

Where *Pi* and *Fi* are the selection and fitness probability of the ith person.

5) Crossover Operation

Crossover operation for GA is a change in the population through producing a new generation which includes some parts driven from parents. On this basis, for the purpose of producing new people, some chromosomes are made through substituting some genes with each other. In this study, two parental chromosomes and bunch's crossing position are determined by random.

6) Mutation Operation

Mutation operation produces some random changes in the structure of the population. In other words, some amounts of chromosomes (weights and biases) change with certain amount of probability. Through using genetic algorithm operators, the network weights and biases are determined properly [16, 17].

DOI 10.5963/JCET0201008

III. RESULTS AND DISCUSSIONS

After structuring network and coding the present parameters through GA, several structures are trained through genetic algorithm. The number of experiments is 6 which are analyzed in networks of 2 neurons to 37 neurons. Table (1) shows the result of these experiments for train, test and validation data. The best number for neurons in each experiment is shown between parentheses in the table. As it is illustrated in the table (1), R-Square for all the experiments is more than 0.98. It means that all the network structures with a hidden layer can predict the used amount of peak of load well. Since the exceeding number of neurons leads to the augmentation of network parameters, selecting a network with the less number of parameters is of high priority. So, the optimized structure with 3 neurons is selected as the best network structure. It needs to be mentioned that the augmentation of the network parameters leads to the augmentation of the overtraining probability in the network which it leads to a dramatic reduction in the efficiency of the network in forecasting.

After approving the ability of the network in forecasting the used peak of load, the graphs related to the amount of peak of loads against days of the year are shown in fig (4) for the whole data. As it is quite obvious in the graph, the mentioned network has successfully evaluated the experimental data for all the four seasons of the year.

It worth mentioning that for this prediction, no supplementary data except for the data from used peak of load has been used which shows a priority over the previous researches with regard to simplicity and preciseness

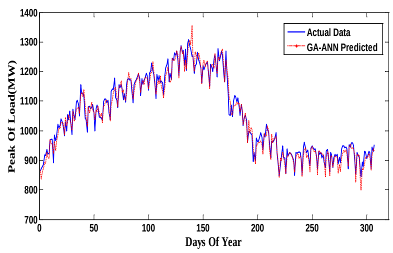

Fig. 4 The comparison between the predicted value and the actual peak value for days of the year

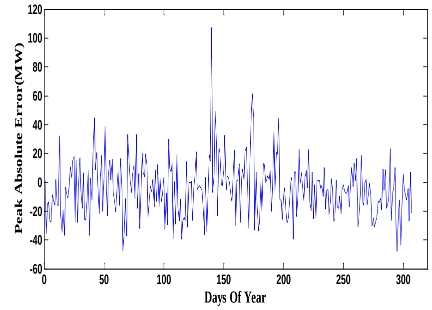

Fig. 5 The difference between predicted value and actual data for the whole year

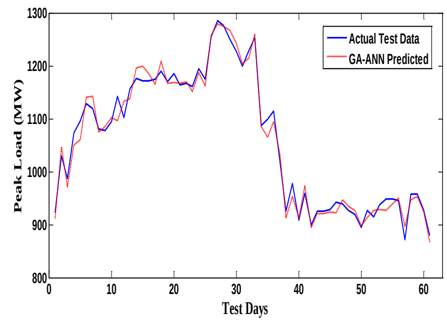

Fig. 6 The comparison between the predicted results and experimental data for test data

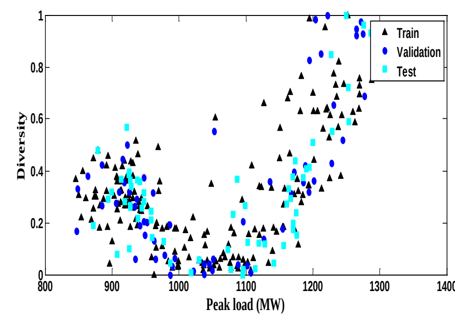

Fig. 7 Comparison of data scattering for train, test and validity data

TABLE (1): THE COMPARISON OF DIFFERENT STRUCTURES BASED ON FORECASTED ERROR

trial	Neuron interval(best)	RMSE(total)	RMSE(train)	RMSE(validation)	RMSE(test)	R-Square(train)	R-Square(validation)	R-Square(test)
1	2-7(3)	21.64	22.63	23.64	15.64	0.9837	0.9837	0.9924
2	8-13(9)	19.03	20.91	21.24	20.75	0.986	0.9871	0.9878
3	14-9(15)	21.51	21.72	18.56	19.83	0.9894	0.9865	0.9867
4	20-5(21)	18.26	24.40	18.55	19.80	0.98968	0.98132	0.98743
5	26-1(31)	19.72	23.92	18.83	20.12	0.98966	0.98381	0.98797
6	32-7(32)	21.27	25.42	17.23	19.93	0.99129	0.98442	0.98874

DOI 10.5963/JCET0201008

For the purpose of analyzing the results of the model, the error graph for whole data is drawn in fig (5). According to this graph, except for two cases, the maximum amount of difference between this model's prediction and experimental results is about 40 MW. It means that in order to guarantee using of this model it is better mention this error in the model as the certainty span.

Therefore, if we consider the certainty span of the model as ±40 MW, the absolute error between the actual data and predicted data will change in this range.

One of the neural networks problems is its disability in good generalization in case of selecting a bad network structure. In this article for survey this doubt, generalization ability of network is tried with test data. The graph of the model's results for test data and actual value are shown in fig (6). As it is illustrated in this figure, the high precision of the model in prediction indicates its high ability generalizing the results. The changing process in the value of peak load in this figure indicates the similar process of changing data and also the presence of the data related to 4 seasons in the test data.

Two most fundamental research issues in analyzing data base are simplicity and diversity of sampling. The issue of diversity can be defined by a diverse subcategory of system's parameters.

For this purpose, for each data base, *N* was considered a parameter resulted from *m* related parameter which any Xi parameter is illustrated in the form of a vector in the equation 11.

$$X_i = (x_{i1}, x_{i2,} x_{i3} \ldots x_{im})^T \quad i = 1, 2, \ldots, n. \tag{11}$$

$$X=(X_1, X_2, X_3, \ldots, X_i)^T; \tag{12}$$

The upper title of T illustrates the transposed matrix vector. The amount of interval between two parameters of X_i and X_j can be calculated through the mean of the Euclid's intervals of each sample in comparison to other samples as in equation 13.

So, Xij defines the amount of jth variable of Xi parameter. The data base resulted X is illustrated in the form of a matrix Xn*m in the equation 12 [18].

$$d_{ij} = || X_i - X_j || = \sqrt{\sum_{k=1}^{m} (X^{ik} - X^{jk})^2} \tag{13}$$

The norm of the interval based on the variants of the parameter is as equation 14.

$$d_i = \frac{\sum_{j=1}^{n} d_{ij}}{n-1} \quad \text{for} \quad i=1, 2, 3, \ldots, n, \tag{14}$$

Then average distance is normalized in [0 1]. Average normalized distance of samples against load consumption peak for the collection of the data shown in figure (7) which shows scattering in train, test and validation data. From the figure we can find out that elements' structure in all 3 collections are acceptably scattered.

Training collection sufficiently proves model endurance and scattering of test and validity collection in order to forecasting.

IV. COMPARING RESULTS WITH PREVIOUS METHODS

Results of 5 models have been shown in table (2) in order to compare the results of this model with previously presented models.

TABLE 2- COMPARISON OF OBTAINED RESULTS WITH SEVERAL REFERENCES

References	MRE%
This article	1.207
2	1.57
19	2.43
20	2.02
21	1.740
22	4.04

Table 2 shows, the best previously presented model is with the MRE% equal to 1.57 for test data. But in this model MRE is equal to 1.207. This model has 23% decrease in comparison with the best previous model and 7o% decrease in comparison with the worst model. These results confirm the used method in this article and superiority of genetic algorithm in finding total minimum in comparison with algorithms based on Newton.

CONCLUSION

This paper presented an integrated genetic algorithm (GA) and artificial neural network (ANN) to estimate and predict consumed load peak. Genetic algorithm has been used in this method which has more power for finding total minimum in comparison with algorithm ba sed on Newton. This method could forecast consumed load peak of west province of IRAN with high precision in comparison with previous models.

Several structures of ANN model with a hidden layer have been checked. The best of them was the one with 3 neutrons because of few parameters and low Mean relative error percent for test data equal to 1.207. Results of structures of two hidden layers have not been considered in this article due to lack of high precision.

Data scattering test was done in order to test correct choosing of test, train and validation data. This matter confirmed correctness of choosing data collection.

For guaranty of using outcomes of the model, we had to consider a certain value of error as the reliable span for model forecasting. After checking the graph of error/span relative reliability of model outcomes was ±40MW in a way that we can claim that real data will be followed.

REFERENCES

[1] M. Peng, N.F. Hubele, and G.G. Karady."Advancement in the application of Neural Networks for Short-Term Load" Forecasting. IEEE Transactions on Power Systems, 7:250–257, 1992

DOI 10.5963/JCET0201008

[2] M. Ghomi, Mahdi Goodarzi, Mahmud Goodarzi.” Peak Load Forecasting of electric Utilities for West province of IRAN by using Neural Network without Weather information” 12th International Conference on Computer Modeling and Simulation. 978-0-7695-4016-0/ UKSIM 2010.

[3] Bahram Hemmateenejad, Mohammad A. Safarpour, Fariba Taghavi, Application of ab initio theory for the prediction of acidity constants of some 1-hydroxy-9,10-anthraquinone derivatives using genetic neural network, Journal of Molecular Structure (Theochem) 635 (2003) 183–190.

[4] T.S. Lia, C.T. Sub, T.L. Chiangc, Applying robust multi-response quality engineering for parameter selection using a novel neural–genetic algorithm, Computers in Industry 50 (2003) 113–122.

[5] Yas Abbas Alsultanny, Musbah M. Aqel, Pattern recognition using multilayer neural-genetic algorithm, Neurocomputing 51 (2003) 237 – 247.

[6] Julio Caballero, Miguel Garrigab and Michael Ferna´ndeza, 2D Autocorrelation modeling of the negative inotropic activity of calcium entry blockers using Bayesian-regularized genetic neural networks, Bioorganic & Medicinal Chemistry 14 (2006) 3330–3340.

[7] Maykel Pe´rez Gonza´lez, Julio Caballero, Alain Tundidor-Camba, Aliuska Morales Helguerab, Michael Ferna´ndez, Modeling of farnesyltransferase inhibition by some thiol and non-thiol peptidomimetic inhibitors using genetic neural networks and RDF approach, Bioorganic & Medicinal Chemistry 14 (2006) 200–213.

[8] Takashi Hiranoa, Jun-Ichi Mogamia, Ryuichi Hiratab, Yasumi Fujinumab, Comparisons of gap-filling methods for carbon flux dataset: A combination of a genetic algorithm and an artificial neural network,Makoto Oobaa, ecological modelling 198 (2006) 473–486.

[9] Integration of artificial neural networks and genetic algorithm to predict electrical energy consumption, A. Azadeh, S.F. Ghaderi, S. Tarverdian, M. Saberi, Applied Mathematics and Computation 186 (2007) 1731–1741.

[10] H.M.Wei, G.H. Su,W.X. Tian, S.Z. Qiu,W.F, Study on the characteristic points of boiling curve by using wavelet analysis and genetic neural network , Ni, Nuclear Engineering and Design xxx (2009) xxx–xxx.

[11] Yong He, Kejun Zhu, Siwei Gao, Ting Liu, Yue Li, Theory and method of genetic-neural optimizing cut-off grade and grade of crude ore, Expert Systems with Applications 36 (2009) 7617–7623.

[12] Huiming Wei, G.H. Su, S.Z. Qiu, Weifeng Ni, Xingbo Yang, Applications of genetic neural network for prediction of critical heat flux, International Journal of Thermal Sciences xxx (2009) 1–10.

[13] Saeid Nahi "Load Forecasting on Special Days & Holidays in Power Distribution substation Using Neural&Fuzzy Networks", International Conference on Computational Intelligence for Modeling Control and Automation, and international Conference on Intelligent Agents. 0-7695-2731-0/06, 2006.

[14] Fang Liu, Raymond D. Findlay, " A neural network based short term electric load forecasting Ontario Canada " International Conference on Intelligent Agents. Web technologies and Internet Commerce, 0-7695-2731-0/06, 2006.

[15] H.S. Hippert, C.E. Pedreira, and R.C. Souza. Neural Networks for Short-Term Load Forecasting: A Review and Evaluation. IEEE Transactions on Power Systems, 16:44–55, 2001.

[16] Randy L. Haupt, Sue Ellen Haupt, practical genetic algorithms, Published by John Wiley & Sons, Inc., Hoboken, New Jersey 2004.

[17] Mitsuo Gen and Runwei Cheng, Genetic Algorithms and Engineering Design, John Wiley & Sons, Inc., 2000.

[18] A.G. Maldonado, J.P. Doucet, M. Petitjean, B.T. Fan, Mol. Divers. 10 (2006) 39.

[19] H.S. Hippert, C.E. Pedreira, and R.C. Souza. Neural Networks for Short-Term Load Forecasting: A Review Transactions on Power Systems and Evaluation. IEEE, 16:44–55, 2001.

[20] Lalit Mohan Saini, Peak load forecasting using Bayesian regularization, Resilient and adaptive back propagation learning based artificial neural networks, Electric Power Systems Research 78 1302–1310 , 2008.

[21] N. Amjady , F.Keynia, Short-term load forecasting of power systems by combination of wavelet transform and neuro-evolutionary algorithm , Energy 34 46–57 , 2009.

[22] E. Gonzalez-Romera, M.A. Jaramillo-Moran, D. Carmona-Fernandez, Monthly electric energy demand forecasting with neural networks and Fourier series, Energy Conversion and Management 49 3135–3142 , 2008.

Mahdi Goodarzi received the B.S. degree in electrical Engineering from Shahid Chamran University Of Kerman in 2006 and the M.S. degrees in electronics from Razi University in 2009. His current research interests include Analog and Digital IC design, Artificial Intelligences, Modeling & Smart Control.

Mahmood Goudarzi received the BS degree in electrical engineering from Islamic Azad University Arak branch,Arak, Iran, in 2002, MS degrees in electrical engineering from Shahrood University of Technology, Shahrood, iran, in 2008 , respectively.

He is currently working as an educator in the Department of Electrical Engineering of Islamic Azad University, Touyserkan branch, Iran and Sama technical and vocational training college, Islamic Azad University, Broujerd Branch, Broujerd,Iran. His research interests include modeling and stability analysis of dynamical systems, electrical mashines, power electronic, applied mathematics, nonlinear control theory, neural network, fuzzy systems, bifurcation theory and its application into power systems stability analysis, and Design and build intelligent robots.

DOI 10.5963/JCET0201008