INTEGRATION OF MULTI LAYER PERCEPTRON AND DESIGN OF EXPERIMENTS FOR FORECASTING HOUSEHOLD ELECTRICITY CONSUMPTION

A. Azadeh , Z.S. Faiz

Department of Industrial Engineering and Research Institute of Energy Management and Planning, Faculty of Engineering, University of Tehran, P.O. Box 11365-4563, Tel: +98 21 6409774, Fax: +98 21 6461024,

[1]aazadeh@ut.ac.ir or ali@azadeh.com

[2] feiz1361@yahoo.com

Abstract

Due to various seasonal and monthly changes in electricity consumption, it is difficult to model it with conventional methods. This paper illustrates an Artificial Neural Network (ANN) approach based on supervised multi layer perceptron (MLP) network for household electricity consumption forecasting. This is the first study which uses MLP for forecasting household electricity consumption. Previous studies base their verification by the difference in error estimation. However this study shows the advantage of MLP methodology through design of experiment (DOE). Moreover, DOE is based on analysis of variance (ANOVA) and Duncan Multiple Range Test (DMLT). Furthermore, actual data is compared with ANN-MLP and conventional regression model. *The significance of this study is integration of MLP and DOE for improved processing, development and testing of household electricity consumption.* Moreover, it would provide more reliable and precise forecasting for policy makers. To show the applicability and superiority of the integrated approach, annual household electricity consumption in Iran from 1974 to 2003 was collected for processing, training and testing purpose.

Keywords:

Artificial Neural Network, Multi Layer Perceptron, Forecasting, DOE, DMRT, ANOVA, Household, Electricity Consumption

1. Introduction

Forecasting electricity consumption is a relatively difficult task. Electricity consumption represents two essential attributes, on one hand it shows the strong annual changes and on the other hand it clearly shows the increasing trend (Figure 1). Furthermore, the time series is affected by other variations that make the problem hard to model. Artificial Neural Networks (ANNs) are the strong rival of regression and time series in forecasting. ANNs are suitable for modeling this kind of problem with unknown factors. The target is to find the essential structure of data to forecast future consumption with less error.

ANNs have been used in nonlinear systems modeling and simulation. One of the most useful and interesting factors of ANNs is forecasting time series. This application of ANNs is suited where static condition or other conditions where using classic techniques are not suitable and applying time series is complicated. It can be also applied to energy forecasting problems. Some of these applications are short and medium term load forecasting [38,3], adaptive price forecasting[18], forecasting transport energy demand [42].Also in some cases ANN can give us a better output if it is trained with the preprocessed data [1,2]

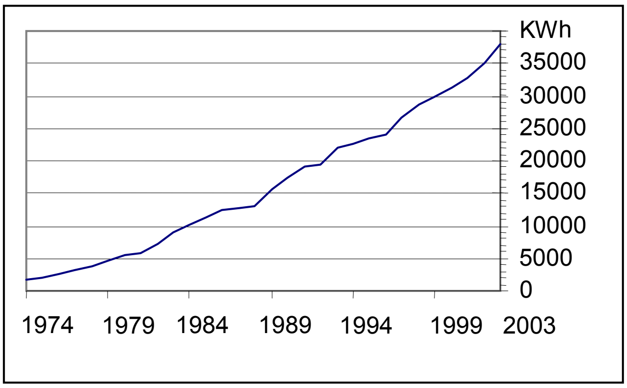

Figure 1: The trend annual energy consumption in Iran from 1974 to 2003

For estimating energy demand, especially electricity, in Iran, we can point to the introductory report of establishing energy model in Iran that Plan and Budget Organization Planning and Budget Organization Estimation of energy carriers' demand in Iran, Technical Report, Operation Research Center, Iran, 1978. has reported for energy carrier's demand. In addition, it has applied econometrics models for forecasting energy supply and demand Planning and Budget Organization

Formulating econometric model for supply and demand, Technical Report, Research Institute of Planning and Development, Iran, 1988. Also Research Institute of Planning and Development Planning and Budget Organization, Estimate demand for energy in Iran, Technical Report, Supreme Research Institute in Planning and Development, Iran, 1982. conducted a study by the name of "energy demand estimation". In this study, the oil product, electricity and natural gas demand models have been estimated for various sectors and total economy. In addition, [26] investigated and estimated electricity demand in residential and industry sectors in Lorestan province of Iran. [27] investigated the stability of demand for energy and its effective factors. [30] forecasted electricity consumption by econometric methods.

In following section ANNs, moving average and data preprocessing method are introduced. Next, the variables for estimating electricity consumption are introduced. The model is then compared with conventional regression model. This will be achieved with the aid of analysis of variance (ANOVA) and Duncan experiment to compare all pairs of treatment means[12].

2. Artificial Neural Network

In general, ANNs are simply mathematical techniques designed to accomplish a variety of tasks. The research in the field has a history of many decades, but after a diminishing interest in the 1970's, a massive growth started in the early 1980's. Today, Neural Networks can be configured in various arrangements to perform a range of tasks including pattern recognition, data mining, classification, forecasting and process modeling [35,5]. ANNs are composed of attributes that lead to perfect solutions in applications where we need to learn a linear or nonlinear mapping. Some of these attributes are: learning ability, generalization, parallel processing and error endurance. These attributes would cause the ANNs solve complex problem methods precisely and flexibly.

ANNs consists of an inter-connection of a number of neurons. There are many varieties of connections under study, however here we will discuss only one type of network which is called the Multi Layer Perceptron (MLP). In this network the data flows forward to the output continuously without any feedback. (Figure 2a and 2b) show the general structure of MLP and a typical three-layer feed forward model used for forecasting purposes. The input nodes are the previous lagged observations while the output provides the forecast for the future value[22,41]. Hidden nodes with appropriate nonlinear transfer functions are used to process the information received by the input nodes. The model can be written as:

$$y_t = \alpha_0 + \sum_{j=1}^{n} \alpha_j f\left(\sum_{i=1}^{m} \beta_{ij} y_{t-i} + \beta_{0j} \right) + \varepsilon_t \qquad (1)$$

Where m is the number of input nodes, n is the number of hidden nodes, f is a sigmoid transfer function such as the

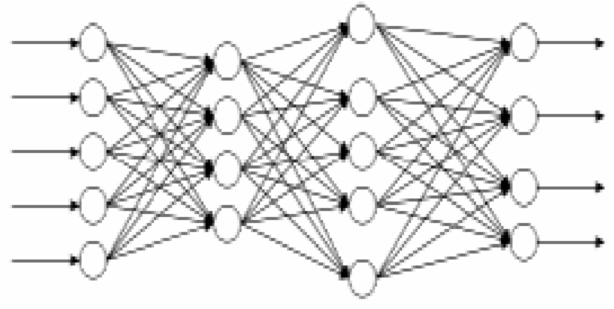

Figure 2a: General structure of the MLP networks

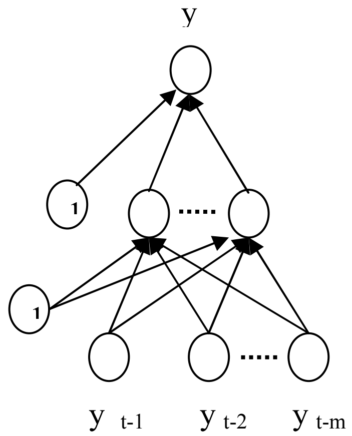

Figure 2b: A three layer MLP network

logistic: $f(x) = \frac{1}{1+\exp(-x)}$. $\{\alpha_j, j = 0, 1, ..., n\}$ is a vector of weights from the hidden to output nodes and $\{\beta_{ij}, i= 1, 2, \ldots, m; j = 0, 1, \ldots, n\}$ are weights from the input to hidden nodes. α_0 and β_{oj} are weights of arcs leading from the bias terms which have values always equal to 1. Note that Equation (1) indicates a linear transfer function is employed in the output node as desired for forecasting problems. The MLP's most popular learning rule is the error back propagation algorithm. Back Propagation learning is a kind of supervised learning introduced by Werbos [35] and later developed by Rumelhart and McClelland [13]. At the beginning of the learning stage all weights in the network are initialized to small random values. The algorithm uses a learning set, which consists of input – desired output pattern pairs. Each input – output pair is obtained by the offline processing of historical data. These pairs are used to adjust the weights in the network to minimize the Sum Squared Error (SSE) which measures the difference between the real and the desired values over all output neurons and all learning patterns[16,24]. After computing SSE, the back propagation step computes the corrections to be applied to the weights.

The ANN models are researched in connection with many power system applications, short-term forecasting being one of the most typical areas. Most of the suggested models use MLP networks see for example [21,40,25,17]. The attraction of MLP has been explained by the ability of the network to learn complex relationships between input and output patterns, which would be difficult to model with conventional algorithmic methods.

There are three steps in solving an ANN problem which are 1) training, 2) generalization and 3) implementation.Training is a process that network learns to recognize present pattern from input data set. We present the network with training examples, which consist of a pattern of activities for the input units together with the desired pattern of activities for the output units. For this reason each ANN uses a set of training rules that define training method.

Generalization or testing evaluates network ability in order to extract a feasible solution when the inputs are unknown to network and are not trained to network. We determine how closely the actual output of the network matches the desired output in new situations. In the learning process the values of interconnection weights are adjusted so that the network produces a better approximation of the desired output. ANNs learn by example. They cannot be programmed to perform a specific task. The examples must be selected carefully otherwise useful time is wasted or even worse the network might be functioning incorrectly. The disadvantage is that because the network finds out how to solve the problem by itself and its operation can be unpredictable. In this paper the effort is made to identify the best fitted network for the desired model according to the characteristics of the problem and ANN features.

3. Model variables

In order to estimate annual electricity household consumption, we introduce electricity household consumption with 5 input variables which are: 1) electricity price, 2) TV price index, 3) refrigerator price index, 4) urban household size and 5) urban household income. The actual data is composed of 30 years of annual household electricity consumption in Iran. Moreover, 22 years of the actual data is used for training the ANN and regression models and the remaining 8 years are to be used for testing and comparing the two models with actual 8 years of data. (Table 1) presents the data for the 5 inputs and electricity energy consumption in Iran from 1974 to 2003. Because of seasonal and monthly trend in electricity consumption it is difficult to forecast its trend by conventional methods such as regression and time series. Therefore, ANNs seems to be ideal for such unknown and fluctuating behavior. In most forecasting problems past data is used as input data to generate output data [38]. This is done as follows:

- x_1: Annual household electricity price;
- x_2: Annual TV price index;
- x_3: Annual refrigerator price index;
- x_4: Annual urban household size ;
- x_5: Annual urban household income;
- $f(x_1, x_2, x_3, x_4, x_5)$: Household electricity consumption demand (KWh).

The first five variables are defined as the input variables and the last one is the output variable. We can write a logarithmic relation as in equation (4a) with b_i for i = 1…5 as constants. The constants will be identified by Least Square technique using the first 22 rows of data given in (Table 1) (1974-1995). Keep in mind that the last 8 rows of data will be saved for tests and experimentations. After applying the data in (Table 1) equation (4b) is obtained showing all variables are important.

$$Ln\, f(x_1, x_2, x_3, x_4, x_5) = b_1 + b_2 Ln\, x_1 + b_3 Ln\, x_2 + b_4 Ln\, x_3 + b_5 Ln\, x_4 + b_6 Ln\, x_5 \quad \textbf{(4a)}$$

$$Ln\, f(x_1, x_2, x_3, x_4, x_5) = -\, 28.17335 - 0.476667 Ln\, x_1 + 0.349910 Ln\, x_2 - 0.293425 Ln\, x_3 + 16.06393 Ln\, x_4 + 0.908539 Ln\, x_5 \quad \textbf{(4b)}$$

Dependent Variable: $Ln\, f(x_1, x_2, x_3, x_4, x_5)$

Method: Least Squares

Sample: 1974-1995

Included Observations: 22

Probability	t-Statistic	Standard Error (SE)	Coefficient	Variable
0.0000	-8.163948	3.450946	-28.17335	b_1
0.0209	-2.518095	0.189297	-0.476667	$Ln\, x_1$
0.0064	3.065511	0.114144	0.349910	$Ln\, x_2$
0.0320	-2.315024	0.126748	-0.293425	$Ln\, x_3$
0.0000	7.831631	2.051160	16.06393	$Ln\, x_4$
0.0000	5.425735	0.167450	0.908539	$Ln\, x_5$
9.184379	Mean dependent var		0.978281	R-squared
0.855716	S.D. dependent var		0.972565	Adjusted R-squared
-0.864139	Akaike info criterion		0.141736	SE of regression
-0.571609	Schwarz criterion		0.381692	Sum squared residual
171.1610	F-statistic		16.80174	Log likelihood
0.000000	Prob (F-statistic)		1.327355	Durbin-Watson stat

Year	Household electricity consumption (KW/h) CE	Electricity Price PE	TV price index CPT	Refrigerator price index CPR	Urban household size BC	Urban household income RC
1974	1620	2.33	2.3	1.85	4.6	319873
1975	2034	2.57	2.23	1.81	4.6	360123
1976	2620	2.8	2.2	1.77	4.7	402857
1977	3238	3.57	2.22	1.92	4.7	448729
1978	3797	3.67	2.28	2.08	4.78	486980
1979	4702	3.1	2.65	2.21	4.8	514446
1980	5479	4.2	2.11	2.43	4.81	608849
1981	5809	4.5	3.64	2.76	4.79	658127
1982	7350	4.2	4.56	3.03	4.83	709601
1983	8857	3.9	5.43	3.48	4.8	918386
1984	10069	4.2	6.82	4.04	4.8	1034170
1985	11316	4.2	8.81	4.26	4.8	1037080
1986	12416	4.2	12.59	7.32	4.8	1126639
1987	12668	4.5	25.99	15.93	4.82	1149296
1988	12994	5.23	36.48	24.7	4.84	1339970
1989	15791	5.27	34.78	29.53	4.86	1467436
1990	17344	5.52	29.63	35.35	4.88	2010547
1991	19128	5.7	26.08	33.42	4.9	2840826
1992	19509	7	23.34	31.14	4.8	3541277
1993	22143	9.86	24.17	33.3	4.8	4425611
1994	22473	11.8	37.99	55.88	4.7	5729732
1995	23374	15.5	63.27	85.49	4.7	7368243
1996	23993	20.45	100.53	101.8	4.7	9879358
1997	26523	26.08	100	100	4.75	12115656
1998	28686	31.2	101.5	105.05	4.72	15151894
1999	29754	41.4	118.6	128.67	4.72	18564952
2000	31266	58.32	118.87	137.22	4.73	22387725
2001	32891	65.11	118.6	139.4	4.75	25831527
2002	34946	72.93	117.4	139.4	4.7	33104868
2003	37967	85.97	114	145.3	4.65	39202427

Table 1: The input indicators versus output indicator or household electricity consumption

4. Running the model

One of the most useful neural networks used in regression analysis is Back Propagation learning algorithm [37,13]. Layered ANNs with only one hidden layer using sigmoid function nodes can closely approximate any continuous function[37.15]. Therefore, only one hidden layer is required to generate an arbitrary function. Several MLP networks were generated and tested. The transfer function for the first layer was linear, for all hidden layers were sigmoid and for the last one was linear. Back propagation algorithm was used to adjust the learning procedure and 22 rows of data selected to test the network. The results of the 4 best models and their errors are shown in (Table 2). These results are derived from 22 rows of unlearned data. Error which estimated by Minimum Absolute Percentage Error (MAPE) is calculated from the following equation:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|EC_{i(estimated)} - EC_{i(actual)}\right|}{EC_{i(actual)}} \qquad (5)$$

Where $EC_{i(estimated)}$ is the estimated household electricity consumption and $EC_{i(actual)}$ is the actual value of household electricity consumption.

MLP Model number	1	2	3	4
Learning method	BP, momentum, weight decay	BP	BP, momentum, weight decay	BP, momentum
Number of neurons in the first hidden layer	10	12	12	12
Relative error	0.0475	0.053	0.0466	0.0351

Table 2: The results of running the 4 MLP

It can be seen that the fourth model with 1 hidden layers and (5-12-1) neurons with momentum learning method would obtain smaller error than other models. The graph of the fourth MLP model versus actual data from 1996 to 2003 (test data) is presented in (Figure 3). As mentioned, these results were derived from 22 rows of unlearned data. The error results of the MLP network are also compared with the conventional regression model with respect to three categories of relative error shown in (Table 3). The reader should note that we have considered 3 categories of relative error which are: 1) mean absolute error (MAE), 2) mean square error (MSE) and mean absolute percentage error (MAPE). (Figure 4) shows the difference between the actual data and conventional regression and the actual data and 4th MLP with respect to the 8 years of test data. In addition, the graphical presentation of MAE, MSE and MAPE for the conventional regression and the 4th MLP are shown in (Figure5).

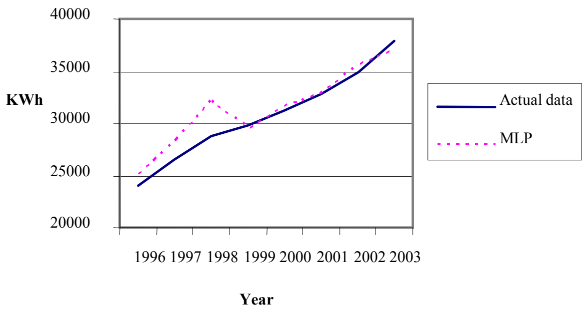

Figure 3: Actual data versus the 4th MLP result

	MAE	MSE	MAPE
MLP errors	0.0108	0.022	0.0351
Regression errors	0.041	0.196	0.117

Table 3: The 4th MLP and regression errors

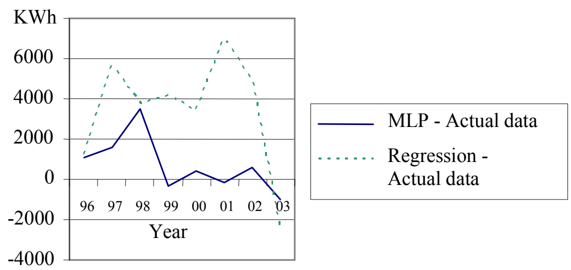

Figure 4: The difference between the actual data and ANN-MLP and regression

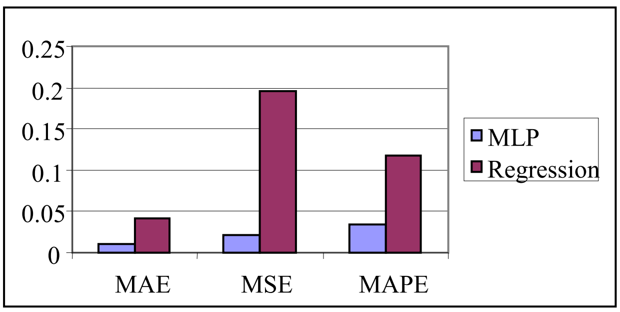

Figure 5: Comparing regression and MLP errors

6. Analysis of variance

The estimated results of the selected ANN, regression method and actual data are compared by analysis of variance (ANOVA). (Table 4) presents the actual, MLP and regression data for the eight years period.

Year	Actual	MLP	Regression
1996	23993	25083.15	25276.79
1997	26523	28146.11	32224.65
1998	28686	32170.78	32441.51
1999	29754	29426.38	33924.81
2000	31266	31651.8	34703.42
2001	32891	32764.41	39912.82
2002	34946	35563.7	39824.37
2003	37967	36963.86	35353.91

Table 4: The eight years test data for actual data, MLP and Regression (KWh)

The experiment was designed such that variability arising from extraneous sources can be systematically controlled. Time is the common source of variability in the experiment that can be systematically controlled through blocking [12].

Therefore a one way blocked design of ANOVA was applied. The results are shown in (Table 5). The test of hypothesis is defined as:

$$\mathbf{H_0: \mu_1 = \mu_2 = \mu_3} \tag{6}$$

$$\mathbf{H_1: \mu_i \neq \mu_j \quad i, j = 1, 2, 3, i \neq j}$$

Where μ_1, μ_2 and μ_3 are the average estimation obtained from actual data, the selected ANN (4th MLP) and regression, respectively. It can be seen from (Table 5) that up to $\alpha = 0.004$ the null hypothesis is rejected because of P-value. Moreover, at $\alpha = 0.01$, we have $F = 8.35$ and $F_{\alpha = 0.01} = 6.51$. Now in order to find which of treatment means (regression or ANN) is closer to actual data, Duncan's Multiple Range Test is applied in the next section.

A . Duncan's Multiple Range Test

In order to perform Duncan's we should find the standard deviation for each treatment mean calculated as:

$$S_{\bar{y}_{i.}} = \sqrt{\frac{MS(error)}{b}} \tag{7}$$

Groups	Count	Sum	Average		
Actual data	8	246026	30753.25		
Neural Network	8	251770.19	31471.27		
Regression	8	273662.28	34207.785		
Source of Variation	Sum Square	Degrees of Freedom	Mean Square	F	F (α = 0.01)
Between Groups (treatments)	53167642	2	26583821	8.35	6.51
Blocks (years)	356467566	7	50923938	16.00	4.28
Within Groups	44555760	14	3182554		P-value is 0.004
Total	454190968	23			

Table 5: The ANOVA table for the regression, actual data and ANN-MLP

Then we should find Rp values that calculated like below:

$$R_p = r_\alpha(p, f) S_{\bar{y}_{i.}} \qquad (8)$$

$r_\alpha(p, f)$ is driven from the Duncan's test table. After sorting the mean treatment, we can compare each treatment as follows:

$\overline{y_{1.}} = 30753.25$, $\overline{y_{2.}} = 31471.27$, $\overline{y_{3.}} = 34207.785$

$s_{\overline{y_{i.}}} = 627.749$,

$r_{0.01}(2,14) = 4.21, r_{0.01}(3,14) = 4.42$

$$R_2 = r_{0.01}(2,14) s_{\bar{y}_{i.}} = 4.21 * 627.749 = 2642.82$$

$$R_3 = r_{0.01}(3,14) s_{\bar{y}_{i.}} = 4.42 * 627.749 = 2774.65$$

Comparing treatments 3 with 1 = 34207.785-30753.25 3454.54 >2774.65 ➔ **$\mu_1 \neq \mu_3$**

Comparing treatments 2 with 1 = 31471.27-30753.25 718.52 <2642.82 ➔ **$\mu_1 = \mu_2$**

We can see that the average of the first (actual data) and second treatment (the selected ANN) are equal at α = 0.01. This shows that the average estimated values of electricity consumption of the selected ANN and actual data are equal at 99% confidence level. Hence, the ANN outputs outperform the conventional regression significantly.

7. Conclusion

This paper focused on ANN and DOE (ANOVA and DMLT) to forecast household electricity consumption. To show the applicability and superiority of the proposed approach actual data of energy consumption in Iran from 1974 to 2003 was used. MLP network was used and applied with the past 22 years of five input variables which are: 1) electricity price, 2) TV price index, 3) refrigerator price index, 4) urban household size and 5) urban household income. After testing all possible networks with 22 rows of unlearned data, we showed that MLP network with instruction of 1 hidden layer and (5-12-1) neurons with momentum learning method had the best output with an error equal to 0.0466 on the test data. The remaining 8 years of data was used as the test data. Moreover, ANOVA was applied to compare the selected ANN, regression and actual data. It was found that at α = 0.01 (with p-value equal to 0.004) the three treatments are not equal and therefore Duncan's Multiple Range Test was used to identify which model is closer to actual data. Moreover, it was shown that the selected MLP has better estimated values for household electricity consumption. This is the first study which uses MLP and time series for forecasting household electricity consumption. Previous studies base their verification by the difference in error estimation. However this study shows the advantage of ANN methodology through analysis of variance (ANOVA). Furthermore, actual data is compared with ANN and conventional regression model. The significance of this study is integration of MLP, time series and ANOVA for improved processing, development and testing of household electricity consumption.

8. Future research

In order to extend the model that has been described in this paper, the seasonal changes from data can be removed and compared with the two models presented in this paper. Also, the results of this study could be further improved by other types of preprocessing methods such as Data Mining, Signal Processing and Genetic Algorithm.

Acknowledgement

This research was partially supported by a grant from the Iranian Ministry of Energy and the Research Institute of Energy Management and Planning at the University of Tehran, Iran.

References

[1]A. Azadeh, S.F. Ghaderi, & S. Sohrabkhani, Improving neural networks output with preprocessed data in electricity consumption forecasting, *Proceedings of The 36th International Conference on Computers and Industrial Engineering*, Taipei, Taiwan, 20-23 June 2006.

[2] A. Azadeh, S.F. Ghaderi, & S. Sohrabkhani, Forecasting electrical consumption by neural network, Proceedings of Energex2006: *Proceedings of the 11th International Energy Conference and Exibition*, Stavanger, Norway, 12-15 June2006.

[3] A. Asar, & J.R. McDonald, A specification of neural network applications in the load forecasting problem, *IEEE transactions on control systems technology*, 2: No. 2, 135-141 (1994).

[4] A. Khotanzad, R.C. Hwang, & A. Abaye, An adaptive modular artificial Neural Network hourly load forecaster and its implementation at electric utilities, *IEEE Transactions on Power Systems*, 10: No. 3, 1716-1722 (1995).

[5] A. Khotanzad, M.H. Davis, A. Abaye, & D.J. Maratukulam, An artificial Neural Network hourly temperature forecaster with applications in load forecasting, *IEEE Transactions on Power Systems*, 11: No. 2, 870-876 (1996).

[6] A. Abraham & B. Nath, A neuro-fuzzy approach for modeling electricity demand in Victoria , *Applied Soft Computing*, 1: No. 2, 127-138 (2001).

[7] A.G. Bakirtzis, J.B. Theocharis, S.J. Kiartzis, & K.J. Satsios, Short term load forecasting using fuzzy neural networks, *IEEE Transactions on Power Systems*, 10: No. 3, 1518-1524 (1995).

[8] B.F. Hobbs, U. Helman, S.Jitprapaikulsarn, S. Konda & D. Maratukulam, Artificial neural networks for short-term energy forecasting: Accuracy and economic value, *Neurocomputing*, 23: 71-84 (1998).

[9] C. Booth, & J.R. McDonald, The use of artificial neural networks for condition monitoring of electrical power transformers, *Neurocomputing*, 23 (1998).

[10] D.E. Rumelhart, & J.L. McClelland (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1, Foundations, MIT Press, Cambridge, MA, 1986.

[11] D.C. Park, M.A. El-Tarawa, R.J. Marks, L.E. Atlas, & M. J. Dasburg, Electric load forecasting using an artificial neural network, *IEEE Transactions on Power Systems*, 6: No. 2, 442-449 (1991).

[12] D.C. Montgomery, *Design & Analysis of Experiments*, Jhon Wiley & Sons, New York, 2001.

[13] D. Rumelhart & J. McClelland, Parallel Distributed Processing: Explorations in the Microstructure of Cogniotion, Vol. 1, Cambridge, MA: The MIT Press (1986).

[14] F.J. Nogales, J. Contreras, & A.J. Conejo, Forecasting next-day electricity prices by time series models, IEEE Tranactions On Power Systems, 17: No. 2 (2002).

[15] G. Cybenko, Approximation by superpositions of sigmoid function, *Mathematics of Control, Signals and Systems*, 2, 303-314 (1989).

[16] G.A. Darbellay, & M. Slama, Forecasting the short-term demand for electricity, Do Neural Networks stand a better chance?. *International Journal of Forecasting* , 16: 71- 83 (2000).

[17] H.S. Hippert, D.W. Bunn & R.C. Souza, Larg neural network for electricity load forecasting: Are they over fitted?, *International Journal of Forecasting*, July-Septamber, 425-434 (2005).

[18] H.Y. Yamina, S.M. Shahidehpourb, & Z. Li, Adaptive short-term electricity price forecasting using artificial Neural Networks in the restructured power markets, *Electrical Power and Energy Systems*, 26: 495-500 (2004).

[19] H.S. Hipperta, D.W. Bunnb, & R.C. Souza, Large neural networks for electricity load forecasting: Are they overfitted?, *International Journal of Forecasting*, 12, 10-20 (2004).

[20] J.V. Ringwood, D. Bofelli & F.T. Murray ,Forecasting electricity demand on short, medium and long time scale using neural networks, *Journal of Intelligent and Robotic System*, (2001).

[21] J.H. Park, Y.M. Park, & K.Y. Lee, Composite modeling for adaptive short term load forecasting, *IEEE Transactions on Power Systems*, 6: No. 2, 450-457 (1991).

[22] J. Abdallah, & A. Al-Thamier, The Economic-environmental neural network model for electrical power dispatching, *Journal of Applied Sciences* (2003).

[23] J. Bao, *Short-term Load Forecasting based on neural network and moving average*, Artificial Intelligence Lab, Dept of Computer Science, 2001.

[24] K.H. Kim, J.K. Park, K.J. Hwang, & S.H. Kim, Implementation of hybrid short-term load forecasting system using artificial neural networks and fuzzy expert systems, *IEEE Transactions on Power Systems*, 10: No. 3, 1534-1539 (1995).

[25] K.Y. Lee, & J.H. Park, Short-term load forecasting using an artificial neural network, *IEEE Transactions on Power Systems*, 7: No. 1, 124-130 (1992).

[26] M. Zamani, *Estimate electricity demand function for the economic sectors*. MS Thesis, Faculty of Economics, University of Tehran, Iran, 1998.

[27] M. Sadeghi, *Demand Stability for energy in Iran*. PhD, Faculty of Economics, University of Tehran, Iran, 1999.

[28] M. Sforna, & F. Proverbio, A Neural Network operator oriented short-term and online load forecasting environment, *Electric Power Systems Research* , 33: 139-149 (1995).

[29] M. Al Mamun & K. Nagasaka, Artifical neural network applied to long-term electricity demand forecasting, *Fourth International Conference on Hybrid Intelligent System* 204-209 (2005).

[30] N. Sadeghi, *Forecasting and modeling electricity demand by an econometric model*. MS Thesis, Faculty of economics, University of Tehran, Iran, 2003.

31) O. Mohammed, D. Park, R. Merchant, T. Dinh, C. Tong, A. J. Azeem Farah, & C. Drake, Practical experiences with an adaptive neural network short-term load forecasting system, *IEEE Transactions on Power Systems*, 10: No. 1, 254-265 (1995).

[32] Planning and Budget Organization *Estimation of energy carriers' demand in Iran*, Technical Report, Operation Research Center, Iran, 1978.

[33] Planning and Budget Organization *Formulating econometric model for supply and demand, Technical Report*, Research Institute of Planning and Development, Iran, 1988.

[34] Planning and Budget Organization, *Estimate demand for energy in Iran*, Technical Report, Supreme Research Institute in Planning and Development, Iran, 1982.

[35] P.I. Werbos, Beyond Regression: new tools for prediction and analysis in the behavior sciences. Ph.D. Thesis, Harvard University, Cambridge, MA, 1974.

[36] R. Beale & T. Jackson, *Neural Computing: An Introduction*, Institute of Physics Publishing, USA, 1998.

[37] S. Wang, Adaptive non-parametric efficiency frontier analysis: A neural-network-based model, *Computers & Operations Research*, 30, 279-295 (2003).

[38] T. Yalcinoz & U.Eminoglu, Short term and medium term power distribution load forecasting by Neural Networks, *Energy Conversion and Management*, 46: 1393- 1405 (2005).

[39] T.M. Peng, N. Hubele & G.F. Karady, Advancement in the application of neural networks for short-term load forecasting, *IEEE Transactions on Power Systems*, 7: No. 1, 250-256 (1992).

[40] T.M. Peng F.N. Hubele, & G. Karady, An adaptive neural network approach to one-week ahead load forecasting, *IEEE Transactions on Power Systems*, 8: No. 3, 1195-1203 (1993).

[41] T.W. Chow & C. T. Leung, Neural network based short-term load forecasting using weather compensation, *IEEE Transactions on Power Systems*, 11: No. 4, 1736-1742 (1996).

[42] Y.S. Murat & H. Ceylan, Use of artificial neural networks for transport energy demand modeling, *Energy Policy*, 2-10 (2004).