LONG TERM LOAD FORECASTING FOR THE EGYPTIAN NETWORK USING ANN AND REGRESSION MODELS

Wagdy MANSOUR
Shoubra Faculty Of Eng. - Benha University- Egypt
wagdy_ibrahim2001@yahoo.com

Mohamed MOENES
Shoubra Faculty Of Eng. - Benha University - Egypt
dr_moenes@yahoo.com

Hassan MAHMOUD
Egyptian Electricity Holding Company - Egypt
hassan04_egypt@yahoo.com

Ahmed GHAREEB
Shoubra Faculty Of Eng. - Benha University - Egypt
Engahmedtg@gmail.com

ABSTRACT

The major concern for every electrical utility is the ability to provide reliable and uninterrupted service to their customers. The challenge becomes more significant with the fast and sharp increasing need for electric energy in the fast developing countries such as Egypt. Load forecasting is mandatory for planning, operation and control of power system. This paper concerns with long term load forecasting and presents a comparison between two models when applied to the Egyptian unified network, these models are Artificial Neural Network (ANN) model and regression model .Data preprocessing techniques have been applied To improve forecasting accuracy of the model. Forecasting capability of each approach is evaluated by calculating two separate statistical evaluations of the Mean Absolute Percentage Error (MAPE) and the Average Absolute Percentage Error (AAVE).

I. INTRODUCTION

Load forecast has been an attractive research topic for many decades and in many countries all over the world, especially in fast developing countries with higher load growth rate.

Many forecasting and statistical methods have already been tried out to solve load forecasting problem with varying degree of success. These methods can be classified into *(Univariate)* such as time series, in which the load is modeled as a function of its past observed values. Previous literature presented various univariate models such as ARMA (Auto-Regressive Moving Average), ARIMA (Auto-Regressive Integrated Moving Average) and ARMAX. The other category is *(Multivariate)* methods, in which the load is modeled as a function of some exogenous factors especially weather and demographic variables. Recently number of studies was presented to examine the influence of different variables on energy consumption [1].

In [2], Artificial Neural Networks (ANN) and regression (Linear and Log-Linear) approaches were for annual electricity load forecasting in Iran. This study presented a model that is affected by two economical parameters which are Real-GDP and Population. Real-GDP has been calculated through dividing Nominal-GDP by CPI (Consumer Price Index) to reflect the effects of inflation. They achieved good forecasts with MAPE of 2.08%.

In [3], presented a comprehensive study for the Egyptian region. This study tested many ANN structures and different load affecting variables such as generated energy, energy sales, GDP, population and average price. It was proved that using econometric variables gives better results.

In the literature to date, short-term demand forecasting has attracted more attention due to its direct impact to power system control, operation and economy. While, long-term forecasting did not receive as much attention, despite their value for system planning and budget allocation. Natures of long term and short term load forecasting are noticeably different. Short term forecasting has less and limited input variables which have similar characteristics unlike long term forecasting parameters that have excessive number, indirect functionality and high uncertainty. From other point of view ANN structure may be simpler for long term hence; its requirements can be provided with good knowledge of the economic and demographic factors. The robustness and accuracy are the most significant requirements for long term load forecasting.

II. ARTIFICIAL NEURAL NETWORK

In this section we provide a brief introduction to ANN (further and deeper knowledge can be found in [4],[5]). Artificial Neural Networks has been motivated right from its inception from the way that the human brain processes information. In general neural networks are simply mathematical techniques designed to accomplish a variety of tasks. ANN is composed of basic computing processing elements (neurons) which connected together in a form of layers. An artificial neuron can be represented by a simple mathematical model which is shown in Figure 1.

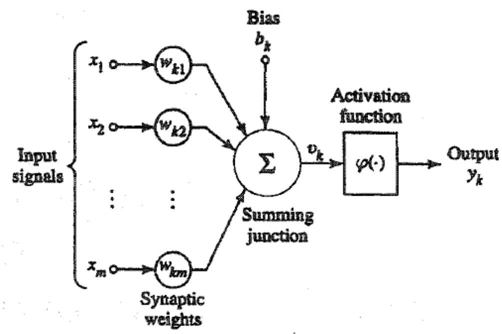

Fig. 1 The mathematical model of the artificial neuron.

The Processing is usually done in two stages: first, the input values are linearly combined and then the result is used as the argument of a nonlinear activation function. The combination uses the weights attributed to each connection, and a constant bias term, represented in the figure by the weight of a connection with a fixed input equal to 1. The activation function must be a non-decreasing and differentiable function; the most common choices are the sigmoidal (s-shaped) functions [6]. An initially constructed ANN is like a newborn child. The neuron weights are initialized with small random numbers when network is first created. A process of learning or training is required to coach this unlearned network by exposure to sample data. The most popular error correction algorithm is the backpropagation rule (BP).So, it will be used in this work.

III. PROPOSED TECHNIQUES

As mentioned before, long term load forecasts will be prepared using ANN model and regression model. ANN model was used to get forecasts utilizing two plans; two years plan and four years plan. Each plan consists of more than two procedures, we get different procedures by changing input variables and methods of preprocessing. While regression model will be used for load forecasting of the next year only. Fig.2 shows a schematic diagram for proposed models.

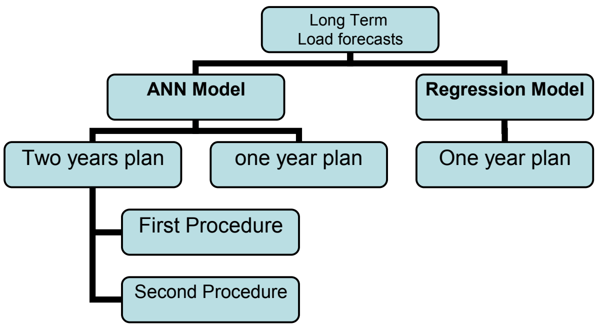

Fig. 2 Proposed Models.

ANN Model

1. Two Year Plan

In this plan, forecasts will be prepared for next two years. Previous literature stated that the most effective parameters on load changes on the long run are Population and GDP.GDP stands for Gross Domestic Product; it is a basic measure of a country's overall economic performance. It is the market value of all final goods and services made within the borders of a country in a year. It is often positively correlated with the standard of living. Fig.3 shows annul peak load, energy sales, Population and GDP in Egypt for 23 years from 1983 to 2006.Note; annual peak load in MW, energy sales in GWh, population in 1000 capita and GDP in Million L.E.

In this plan we scale all the training set data between 0 and 1. To meet this purpose we divide each data set by its norm (i.e. the largest singular value).

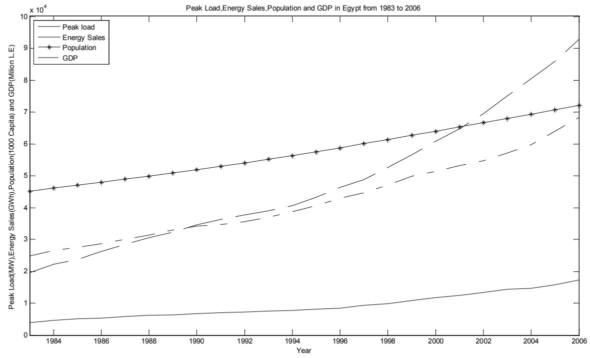

Fig. 3 Peak load,energy sales,population and GDP in Egypt.

1.1 First Procedure

In this procedure ANN's are trained with historical data from year 1993 to year 2004 and designed to predict the peak load for years 2005, 2006. Training data is divided into two sets input set and target set, the target set contains historical annual peak loads for years 1993 to 2004 (twelve values) and the input set is entered to the network as a matrix of 12×4 elements the first column is year index, the second column is the GDP, the third column is the squared values of GDP and the last column is the population of Egypt for the specified period. The choice of these factors is based on the results gained from the correlation analysis done by SPSS. The correlation coefficient between peak load and GDP,GDP^2, POP are 0.994 ,0.994 and 0.982 respectively.

This correlation analysis was done by SPSS (Statistical Package for Social Sciences); it is a computer program used for statistical analysis. Before 2009 it was called SPSS, but in 2009 it was re-branded as PASW (Predictive Analytics Software). SPSS is one of the most powerful and widely used programs for statistical analysis in social science.

Table (1) Various ANN structures, results and corresponding error values for this procedure.

No of neurons / layer	Transfer Function	Training Function	%E	AAVE	
[25,15,8,1]	[tansig,tansig, tansig,purlin]	trainlm	1.7633	3.4102	3.0873
[5,1]	[tansig,purlin]	trainrp	0.7236	0.0918	2.4711
[8,1]	[tansig,purlin]	traingdx	0.4139	0.0186	1.4448
[3,1]	[tansig,purlin]	traingda	1.5809	3.6903	2.1852
[7,1]	**[tansig,purlin]**	**trainrp**	1.6851	1.7739	4.3179
			1.0275	2.1644	1.6357
			0.2104	**0.7531**	**0.0495**
[8,3,1]	[tansig,logsig, purlin]	trainrp	1.2492	0.2222	4.2071
			1.2316	0.0298	4.3222
[6,2,1]	[tansig,purlin, purlin]	Traincgb	1.5526	1.2921	4.2935
[6,2,1]	[tansig,purlin, purlin]	trainrp	0.5753	0.7000	1.3872
			0.7239	0.4586	2.1342
			0.3480	1.1776	0.1446

1.2 Second Procedure

In this procedure, the training set contains years and the multiplication of GDP by population only. Investigation of correlation between peak load and many other variables leads to the fact that if we combine two variables or more we get a new variable which has higher correlation factor to peak load than each individual factor. So when the gross domestic product is multiplied by population a third factor is gained, this variable is just a training variable for the network. Correlation factors between peak load and GDP, population and GDP×population are 0.994, 0.982 and 0.997 so; the third factor was chosen to be in the training set [7].

Table 2. Structures & results for the second procedure.

Neurons / layer	Transfer Function	Training function	%E	AAVE	
[8,1]	**[tansig,purlin**	**trainrp**	**0.186**	**0.20**	**0.47**
[50,10,1]	[purlin,tansig, purlin]	trainrp	0.28	0.02	0.98
			1.98	6.56	0.94
[4,1]	[tansig,purlin]	trainrp	2.14	0.81	6.81
[8,3,1]	[purlin,tansig &purlin]	traincgb	0.758	1.56	1.23
			2.150	7.65	0.54
[8,3,1]	[purlin,tansig &purlin]	trainrp	1.503	3.60	1.99
			1.437	1.14	4.02
			1.564	5.90	0.08
			1.813	2.06	4.50
[6,4,2,1]	[purlin,tansig, tansig&purlin]	traincgb	1.09	0.17	3.71
[8,5,3,1]	[purlin,tansig, tansig&purlin]	trainrp	0.610	1.12	1.12
			0.658	0.05	2.27
[6,1]	[purlin,purlin]	trainrp	1.103	4.06	0.15
			1.355	2.61	2.37
[6,1]	[tansig,tansig]	trainrp	1.186	0.34	3.86
[6,1]	[tansig,purlin]	trainrp	0.722	2.70	0.05
			1.045	0.58	3.14

2. One Year Plan

In this procedure ANN's are designed to forecast both energy sales& peak load for year 2007 and the network is trained with historical data from year 1993 to 2006 (14 observations), training set consists of year index , GDP, GDP*Population and Square of Population . The high correlation factors answer the question why specifically these factors are chosen to train the network among many other parameters. Pre-processing of training data was done By Mapminmax or Mapstd. Both functions are included in MATLAB7.5. Mapminmax preprocesses data so that min. & max. are pre-set values, the default values are -1&1.While, Mapstd processes variables by transforming the mean and standard deviation to pre-set values, default values are 0, 1 respectively. With the same methodology, different ANN structures with different transfer, training, pre-processing functions and different number of neurons are tested. Sometimes various results can be gained for the same structure because initial weights and biases of ANN are varied for each trial. This procedure has several satisfied results obtained by applying the two types of pre-processing to compare between them. Table 3 shows results when the preprocessing function is |Mapminmax. Table 4 shows results when data is preprocessed using Mapstd.

Table (3) Pre-processing function is Mapminmax.

no of neurons/ layer	Transfer function	Training function	MAPE	AAVE	
[10,2]	[tansig,purlin]	Trainlm	1.563	2.678	2.533
[25,8,2]	[putlin,tansig,purlin]	Trainlm	1.026	2.053	1.373
[70,25,2]	[putlin,tansig,purlin]	Trainlm	0.384	0.920	0.365
[35,10,2]	[putlin,tansig,purlin]	Trainlm	1.327	2.723	1.710
			1.454	0.896	3.922
			1.321	3.080	1.339
[12,5,3,2]	[putlin,logsig,tansig,purlin]	Trainlm	0.998	0.700	2.610
[15,10,5,2]	[putlin,logsig,tansig,purlin]	Trainlm	1.110	1.796	1.902
[12,8,3,2]	[putlin,tansig,tansig,purlin]	Traincgb	1.541	3.305	1.845
			2.470	0.140	8.023
			0.659	0.672	1.518
[60,30,2]	[putlin,tansig,purlin]	Traincgb	1.277	4.116	0.177
[7,5,3,2]	[putlin,logsig,tansig,purlin]	Traincgb	0.663	0.973	1.234
			0.453	0.787	0.724
			0.705	1.941	0.424
[7,5,3,2]	[putlin,tansig,tansig,purlin]	Traincgb	1.113	2.049	1.664
			0.373	0.500	0.742
			0.429	1.387	0.055
[12,8,3,2]	**[purlin,tansig,tansig,purlin]**	**Trainrp**	**0.087**	**0.199**	**0.092**
[15,5,2]	[putlin,tansig,purlin]	Trainrp	0.493	0.915	0.729
			1.172	3.570	0.365
			0.633	0.578	1.524
[25,10,2]	[putlin,tansig,purlin]	Trainrp	1.550	4.126	1.070
[12,5,2]	[putlin,tansig,purlin]	Trainrp	0.128	0.410	0.018
			0.319	0.911	0.159
[40,20,2]	[putlin,tansig,purlin]	Trainrp	1.004	2.613	0.749
[7,5,3,2]	[putlin,tansig,tansig,purlin]	Trainrp	0.199	0.595	0.074
			0.251	0.486	0.353

Table (4) Pre-processing function is Mapstd.

no of neurons/ layer	Transfer function	Training function	MAPE	AAVE	
[10,5,3,2]	[purlin,tansig,tansig,purlin]	trainlm	1.255	0.185	3.962
[7,5,3,2]	[purlin,tansig,tansig,purlin]	trainlm	0.546	1.103	0.720
[70,25,2]	[purlin,tansig,purlin]	trainlm	0.287	0.527	0.430
			0.689	1.197	1.101
[35,10,2]	[purlin,tansig,purlin]	trainlm	0.864	1.432	1.448
			0.708	2.062	0.313
[30,15,2]	[tansig,tansig,purlin]	trainlm	0.303	0.332	0.676
[30,15,2]	[logsig,tansig,purlin]	trainlm	0.774	1.246	1.333
			0.581	1.547	0.402
[15,10,5,2]	[putlin,tansig,tansig,purlin]	Trainlm	2.687	8.423	0.607
			1.172	1.081	2.812
[20,10,2]	[tansig,tansig,purlin]	Trainlm	1.555	3.053	2.140
			1.053	2.581	0.944
			0.461	0.162	1.363
[12,8,3,2]	[purlin,tansig,tansig,purlin]	trainrp	2.212	2.448	4.904
			1.659	2.946	2.587
			0.716	0.013	2.354
[10,2]	**[tansig,purlin]**	**trainrp**	**0.284**	**0.074**	**0.865**
			0.627	2.107	0.003
[8,3,2]	[purlin,tansig,purlin]	trainrp	0.843	2.828	0.007
[40,20,2]	[purlin,tansig,purlin]	trainrp	1.804	5.917	0.152
[5,3,2]	[tansig,logsig,purlin]	Trainrp	0.752	1.712	0.804
			0.724	1.037	1.374
			0.746	0.260	2.209
[30,15,10,2]	[purlin,logsig,tansig,purlin]	trainrp	1.300	0.720	3.589
[25,10,2]	[purlin,tansig,purlin]	traingdx	0.509	0.983	0.715
[7,5,3,2]	[putlin,tansig,tansig,purlin]	Traincgb	1.326	3.104	1.335
			1.010	0.083	3.255
[15,8,5,2]	[purlin,tansig,tansig,purlin]	traincgb	1.535	4.675	0.482
			1.423	4.424	0.357
[5,3,2]	[tansig,logsig,purlin]	Traincgb	1.702	4.766	0.944

Regression Model

In this section we will use Linear Regression model to forecast peak load of year 2006.Data used to prepare this model was historical data of years 1983-2005. This model can be summarized as the following:

Forecasted load = C(1)+C(2)*GDP+C(3)*POP

The forecasted load will be calculated as a function of GDP and Population, previous correlation analysis has shown that annual peak load of Egypt depends greatly on both factors GDP and population. So, these factors will be used again in our regression model GDP and Population are called independent variables while the load will be the dependent variable After calculating regression Coefficients(C(1),C(2) and C(3)) the equation will be:

Forecasted load

= (4356.55) + (0.482)*GDP – (0.27)*POP

This equation will be applied for all data contained in data set, and load for each year will be calculated. Figure 4 shows forecasted results from the proposed model and the actual loads.

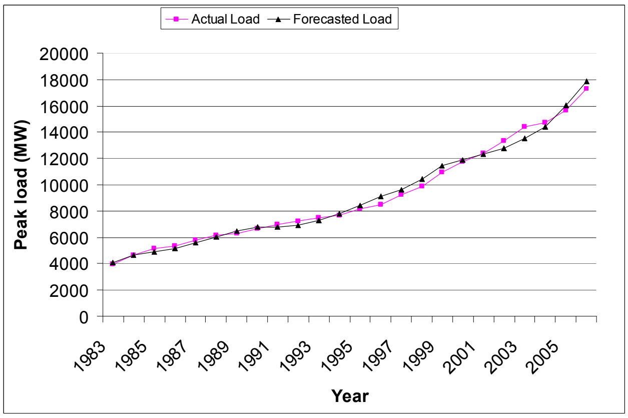

Fig. 4 Actual & Forecasted loads for years (1983-2006).

IV. RESULTS & COMMENTS

1. For ANN model using GDP,GDP an population we get accurate forecasts with MAPE 0.2104% with very simple two layer structure. The first layer contains 7 neurons and only one neuron in the second layer. Training function was Trainrp and transfer functions are Tansigmoid & purlin.

.

2. By changing the training set to use the new factor represented in GDP×POP, we get more reduced error of 0.18%.The used structure consists of two layers with 8 and 1 neuron/layer respectively. The training function is trainrp.

3. By applying this methodology to forecast both energy sales and peak load for year 2007.We get forecasts with MAPE of 0.284 for mapstd preprocessing and 0.087 for mapminmax preprocessing. Experiments have shown that preprocessing using mapminmax achieves more effective training and precise predictions for the same structures.

4. Two variable regression model was tested to forecast peak load of year 2006. This model gives relatively good results, if compared to other regression models proposed in literature. The forecasted load was 17837.8 MW, with percentage error 3.109%, the average error for 24 observations equals 3.278 %.

V. CONCLUSION

The proposed method was utilized to predict peak load and energy sales for Egyptian unified network. Its obvious that the new methodology of combining one or two variables to get a new training variable reduces significantly the forecasting error. Also, preprocessing of training data set has a noticeable effect in improving forecasting accuracy. Two variable regression model was tested to forecast annual peak load. It provides relatively good results but it still not accurate as ANN predictions. ANNs have shown great and unique ability with dealing with nonlinear problems such as load forecasting problem. It overcomes many of regression model drawbacks. It isn't novel to say that but it's a matter of confirmation.

REFRENCES

[1] Henrique Steinherz Hippert, Carlos Eduardo Pedreira, and Reinaldo Castro Souza, 2001, " Neural networks for short-term load forecasting:a review and evaluation"*IEEE Trans. Power Systems*, vol. 16, no. 1, 44–55.

[2] A. Ghanbari, A. Naghavi, S.F. Ghaderi, M. Sabaghian, 2009, "Artificial Neural Networks and Regression Approaches Comparison for Forecasting Iran's Annual Electricity Load", *International Conference of Power Engineering, Energy and Electrical Drives, POWERENG '09,* 675-679.

[3] Hend Abd-El-Monem, 2008, "Artificial intelligence applications for load forecasting," *Ph.D. thesis, Electric power and machines department, Ain-Shams University, Egypt.*

[4] Simon Haykin, 1998, *Neural networks: a comprehensive foundation,* 2nd edition, Prentice Hall Inc., NJ, USA.

[5] Jacek M. Zurada, 1992, *Introduction to artificial neural systems,* West publishing company, MN., USA.

[6] Kittipong Methaprayoon, 2003, "Neural network-based short term load forecasting for unit commitment scheduling", M.sc's thesis, *University of Texas at Arlington, USA.*

[7] Ahmed T. Ghareeb, 2010, "Load Forecasting and system upgrades with consideration of weather parameters", M.sc's thesis, *Benha University, Egypt.*

[8] Rob J. Hyndman and Shu Fan, 2010, "Density Forecasting for Long-Term Peak Electricity Demand", *IEEE Transactions on Power Systems,* vol. 25, no. 2, 1142–1153.

[9] A. D. Papalexopoulos, S. Hao, and T. M. Peng, 1994, "An implementation of a neural network based load forecasting model for the EMS", *IEEE Trans.Power Systems*, vol. 9, no. 4,1956–1962.

[10] Jeffrey M. Wooldridge, 2008, *Introductory Econometrics: A modern approach,* 3rd edition.

[11] Alexander Bonner, 2009, "*Forecasting Models for the German Office Market*",1st edition, GWV Fachverlage GmbH, Wiesbaden, Germany.

[12] Elvira N. Loredo, 2001, "Annual electrical peak load forecasting methods with measures of prediction error", *Ph.D. thesis, Arizona state University,USA.*

[13] O. Mohammed, D. Park, R. Merchant, T. Dinh, C. Tong, A. Azeem, J.Farah, and C. Drake, 1995, "Practical experiences with an adaptive neural network short-term load forecasting system", *IEEE Trans. Power Systems,*vol. 10, no. 1, 254–265.

[14] A. D. Papalexopoulos, and Timothy C. Hesterberg, 1990, "A Regression-based approach to short-term system load forecasting", *IEEE Trans. Power Systems,* vol. 5, no. 4, 1535–1550.

[15] Lidong Duan, Dongxiao Niu and Zhihong Gu, 2008, " Long and medium term power load forecasting with multi-level recursive regression analysis," *Second International Symposium on Intelligent Information Technology Application (IITA),IEEE conferences,* 514-518.

[16] Samprit Chattefuee and Ali S. Had1, 2006, "Regression analysis by example", *4th edition, John Wiley & Sons, Inc*.

[17] Robert F. Engle , Chowdhury Mustafa and John Rice, 1992, " Modeling peak electricity demand", *Journal of forecasting*,Vol.11, 241-251.