APPLICATIONS ON MEDIUM-TERM FORECASTING FOR LOADS AND ENERGY SCALES BY USING ARTIFICIAL NEURAL NETWORK

H. A.E. SALAMA
The Supreme Council of antiquities - Egypt
h.d.61@hotmail.com

A. F. A. EL-GAWAD
Zagazig university, Faculty of Eng. -Egypt
amgawad2001@yahoo.com

S. M. SAKR
Ain Shams University, Faculty of Eng. - Egypt
sohiersakr@yahoo.com

E. A. MOHAMED
Ain Shams University, Faculty of Eng. - Egypt
elsayed_abd_elaleem@yahoo.com

H. M. MAHMOUD
Egyptian Elec. Holding Company- Egypt
hassan01_egypt@lycos.com

ABSTRACT

In this paper, forecasting estimation for medium- term period by using Artificial Neural Network (ANN) techniques has been implemented. The actual data are obtained from The Egyptian Electric Holding Company (EEHC). The study focused on unexpected peaks on working days of summer seasons that have the dangerous effect on safety of electric power system. Medium-term load forecasting is established by two plans, One-year plan and Five-years plan with different designed techniques. The input consists of many parameters such as past annual loads, population, oil price and so on. The outputs are annual loads or the energy sales or both. The results of these techniques are compared with output of the models calculated by EEPU. So, valuable conclusions and recommendations are obtained.

INTRODUCTION

Forecasting is the future planning of the Electric Power Utilities. Development of publics is based on the future planning corresponding to the civilization and the resources of the countries. In Egypt, new sectors have to be developed such as tourism sector beside the conventional sectors such as industry, agriculture and commercial sectors. The Supreme Council of antiquities has recently established the restoration of ancient monuments with the developments of the surrounding areas. This feature is attracting huge number of tourists, so, this requires the suggestion of the future planes of these projects in additional to the tourists villages in famous beaches of Hurghada and Sharm El-Shekh. Then the load forecasting system may have the significant role sharing in the future planning complete the work. In Egypt, Wednesday and Friday are considered the well-known working day and week-end day, respectively, and they are the two categories that have been predicted on the summer season of the year. Unexpected peaks on Wednesdays of summer seasons must be studied well to offer the security of electric power system. In this study, One-year plan and Five-year plan are prepared as a pattern of medium-term forecasting for the peak loads that occur frequently in Wednesdays of August. Special car must be taken for the unexpected change for that day between some successive two years 1992, 1993 and 1995, 1996 and 2002, 2003 respectively. It is deduced that few researches are introduced for medium term load forecasting by using ANN. This may return to the less of available input data for processing. However, the two plans, with different ANNs techniques are evaluated with successfully results. These results are achieved by using Feed forward back propagation under Matlab6.0 toolbox with different transfer functions, Training Functions and preprocessing for the input and output data. Finally, good comparison was made to weigh up the results of these techniques with output of the models calculated by EEHC.

PROBLEM DESCRIPTION

Available data is the real data of electric power system in Egypt. Abnormal and rapid change of annual load curve especially in previous ten years ago became unsuitable data to be used in the predicted future annual load curve. From tedious investigations, load curve, before the year 1990, has a trend of constant growth rate but the next year has a different load curve that would be varying along the second year. Thus, historical data before 1990 must be taken with fear in predicting studies because it may produce unsuitable forecasting outputs. So, shortage of the input data required for ANN construction that helps the network to learn well is noted and may cause incorrect results. Also, annual required demand increases with unexpected growth rate that make sure that there are other parameters that have the impact on the demand. The parameters prepared for prediction can be summarized as following**:** Past annual loads, Average annual temperature, Gross national product, Electricity price, Population, Inflation, Amount of energy, Consumption of energy / household, Oil price, Country development, and Gross domestic product [1], [2]. Regarding to the data of the power system in Egypt, only five parameters would be available as shown later while the others are not permitted [3]**.** In medium-term forecasting, proposed techniques is designed to provide the forecasted output that will be the value of annual load only or annual energy sales only or load and energy sales together. Also, the forecasting output will be for the next year or for the next alternative five years. The following equation estimates the difference between the forecasted and actual load as: Average error % =

$$\frac{1}{N}\sum_{i=1}^{i=N} \frac{\textit{the foreacasted.load} - \textit{the acual.load}}{\textit{the actual.load}} \times 100$$

Where, N is number of observations (used years) of the network. So, the proposed techniques of ANNs reserve their validity and implement satisfied results when the percentage average error will be equal zero or minimum. Proposed techniques can be explained at the following sections.

ONE YEAR PLAN

This technique can be divided into four models as presented in table (1). The network is designed to predict the electric loads or the energy sales or both of them. Electric Power System in Egypt for the next year and consequently the forecasted result can be used as actual data to feed ANN to retrain and predict the output for another next year by using its predicted input vectors of the network. Interval of actual data and forecasted targets of proposed models is based on the investigations mentioned before. Annual load curve varied actually from the year 1990 which its trend became unstable and this may influence on the output results. Then, to examine annual load curves before and after 1990, it requires the change of models interval as mentioned.

Table (1): Input / Output of One Year Plan

Training (P)	**I / P**	1-No. of years 2-E S 3-G D P 4-P O P 5-Elect. P O P 6-L F 7-G D P for Capita.			
	O /P	**First Model**	**Second Model**	**Third Model**	**Fourth Model**
		annual loads from the year 1990 to 2002	annual loads from the year 1981 to 1995	annual energy sales from the year 1981 to 1992	annual loads and energy sales from the year 1981 to 1992
Testing (Pnew)	**I / P**	Same input vectors of P with predicted values for lead time period.			
	O /P	**First Model**	**Second Model**	**Third Model**	**Fourth Model**
		Loads of the year 2003	loads of the year 1996	energy sales of the year 1993	Loads and energy sales of the year 1993.

Where: G D P is the gross domestic product, E S is the energy sales, P O P is the population, Elect. P O P is the electricity population (for person), L F is load factor, and G D P per Capita is the gross domestic product for the person.

ANN has two matrices, P and Pnew. The first matrix is called simulated matrix which contains the historical actual data with certain number of observation (number of used years) to simulate the actual target of ANN and the second Pnew is called testing matrix that has the predicted values of input parameters for the forecasted interval to determine the forecasted output. It is evident to mention that, generally, the inputs of the testing ANN may be predicted or guessed and this may yields to inaccurate outputs. So, accurate programs must be used to determine these values as the forecasted targets are based on them. In this study it is required only to measure the validity of ANN application within odd load curve than other estimating methods, so, the used input and outputs for P and Pnew are already known for the historical data between 1981 &2005. Quality performance of the required goal of P is examined. Therefore, when the performance goal is met, the testing matrix will certainly succeed. The input parameters are absolutely different in values and it may corrupt the system training, these data must be normalized by using ANN generalization Preprocessing. Preprocessing is scaling the inputs and the targets of the network to fall within the specified range. So, the quality performance of P is performed and then, the outputs of Pnew are improved. Scaling the inputs and the targets is divided into two types; Premnmx and Mean & Standard Deviation. Premnmx normalizes the inputs and the targets to fall within the value [-1, 1] while Mean & Standard Deviation normalizes the inputs and the targets so that they will have zero mean and unity standard deviation. Various constructions of ANN are presented by the training. different transfer functions are illustrated and more than one output will be satisfied. It can be deduced that weights of network has a specified initial value for each trial through training program.

Forth model will be explained in some details as an example for the proposed technique.

Characteristic and Structure of Fourth Model ANN

The network of this model is designed to predict two outputs that are the energy sales and load as indicated in table (1). Historical input data are between the years 1981 and 1992 and so, number of observation is 12. The change of ANN construction must be done to perform the network and improve the outputs.

Training and Results of Fourth Model

Several structure of ANN is illustrated with two outputs (energy sales and peak loads) that produced several satisfied results by using two choices of ANN Preprocessing of inputs and outputs. These results include the annual least error of load or forecasting energy sales. So, training after the two preprocessing functions Premnmx and Mean & Standard Deviation may give different results. In fact training functions trained with various initial weights and bias of the network. So, more than one satisfied result for one structure will be produced. All trials provide high quality performance where its required goal is met and then the testing network produces accurate results. The best results of the previous trainings with preprocessing the data can be indicated by the table (2).

Table (2): Results of Various ANN Structure of Fourth Model

No. of neurons, layers	Transfer functions	Training Function and preprocessing	Error %
[30-15-2]	[logsig,tansig,purelin]	Traingdx*	0.128 &0.55
[30-15-2]	**[tansig,tansig,purelin]**	**Trainlm***	**0.159 &0.044**
[20-10-2]	[tansig,tansig,purelin]	Traingdx**	0.536 &0.358
[30-2]	[tansig,purlin]	Trainrp**	0.413& 0.61
[20-10-2]	[tansig,tansig,purlin]	Trainlm**	0.515&0.551

* Premnmx ** Mean & Standard Deviation

It is clear that fourth model has two outputs: forecasted loads and energy sales. The structure with the least error is marked

by bold i.e. **0**.159 and 0.044 for energy sales and peak load when the structure provide the following: Training function is Trainlm, Preprocessing type is Mean & Standard Deviation, Number of neurons are [30-15-2] with three hidden layers and transfer functions are tangent sigmoid and linear purelin. In general performance goal is met of all trainings with various structures of ANN for the four models. Best results gained by ANN construction depends on logistic and tangent sigmoid transfer function of hidden layers and linear transfer function of output layer. Network design achieves successful simulation network with the Trainrp training function. Whatever, the least error of two outputs is provided by more than one structure indicated in training tables. Trainrp has poor quality performance to use in this model. In general, Premnmx preprocessing of ANN inputs gave results that are preferable than Mean & Standard Deviation.

Five Years plan

The annual forecasted load for the sequent five years has been calculated by another technique. Two main models are included in this technique that is explained as follows:

- For the first model it will be divided into two separate networks that have the following characteristics: First network has one target which is the load values. Second network has two targets which are energy sales and load values.
- For the second model, two parameters added to ANN inputs, the multiplication of the population and the energy sales, in addition to the growth rate of annual load. Thus, number of observation and number of parameters are increased in second network of the second technique to improve the final outputs.

Table (3): Input / Output of Five Years Plan Technique

		First Model		Second Model
Training (P)	I / P	1-No. of years 2-E S 3-G D P 4-P O P 5-Elect. P O P 6-L F 7-G D P for Capita 8-G.R.		1-No. of years 2-E S 3-G D P 4-P O P 5-Elect. P O P 6-L F 7-G D P for Capita 8-G.R. 9-E S × P O P.
	O /P	**First Model**		**Second Model**
		Annual loads from the year 1981 to 1992.	Energy sales and loads from the year 1981 to 1992.	Annual loads from the year 1982 to 2000.
Testing (Pnew)	I / P	Same input vectors of P with predicted values for lead time period.		
	O /P	**First Model**		**Second Model**
		Loads for the next 5 years after 1992.	Loads and energy sales for the next five years after 1992.	Loads for the next 5 years after 2000.

Where: G R is the annual growth rate and E S × P O P is the energy sales multiplied by population.

Training of First Network of First Model

The network has 12 observations (number of years) and eight input vectors. They are considered few input number for the network to learn well. The least error of this model that equals 1.45 %, gained by Mean & Standard Deviation preprocessing. Three hidden layers [30,15,1] are provided and the training function is traingdm while Tangent sigmoid is the preferable transfer function

Training of Second Network of First Model

Best result was achieved with average error of 2.38 % and 2.49 %.for energy sales and loads, respectively. It was gained by Mean & Standard Deviation preprocessing. Three hidden layers [40, 15, 2] are provided and the training function is trainlm while Log - Tangent sigmoids were the preferable transfer functions.

Table (4): Actual and Forecasted Values of Peak loads and Energy Sales

Years Index	Forecasted Loads	Actual Loads	Forecasted Energy Sales	Actual Energy Sales
1993	7488.2	7503	39322	38988
1994	7935.3	7675	41792	40577
1995	8260.3	8149	45300	43255
1996	8316.3	8491	47621	46281
1997	8941.6	9235	48915	49139

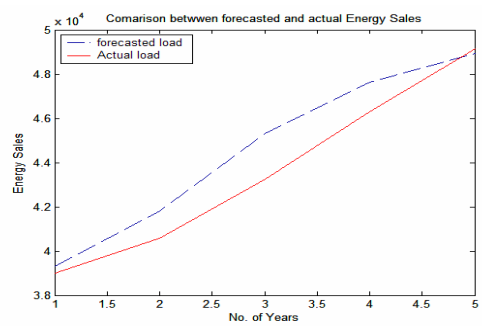

Fig.(1): Comparison Between Forecasted and Actual Energy Sales of Years (1993-1997)

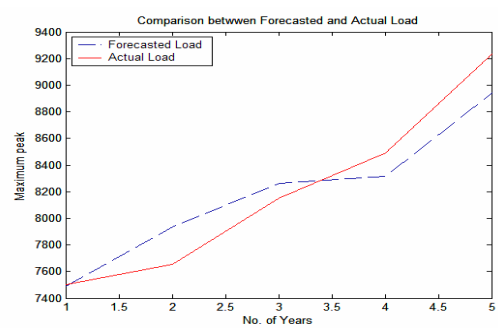

Fig.(2): Comparison Between Forecasted and Actual Loads of Years (1993-1997)

Actual and forecasted energy sales and loads indicated in table (4) and the comparison between actual and forecasted results can be presented by the Figs.(1) and (2), respectively. Sometimes, if energy sales achieve the least error, load values have high percentage error, so, it is preferable to separate them in to different networks to get better result.

Training of Second Model

Various structures of ANN are provided and trained as shown in table (5). The least error is gained by the first structure of the table. It is worth to mention that more input parameters must be added to increase number of used observations to improve the network results. Proposed ANN output, the output calculated by EEHC models and the calculated errors can be shown in table (6).

Table (5): Second Model Training (2001-2005)

No. of neurons, layers	Transfer functions	Training function	Average Error %
[40,10,1]	**[logsig,tansig,purelin]**	**Trainlm****	**6.81**
[60,30,5 1]	[log,tan,tan,pure]	Traingdx*	8.67
[40,15 1]	[tansig,tansig,purelin]	Traingdm** mc=0.85	7.04

* Premnmx ** Mean & Standard Deviation

Table (6): EEHC Outputs, Forecasted Outputs and Actual Load of Second Network

Years Index	EEHC Models	Proposed Techniques	Actual Load	Error of EEHC Tech.	Error of Proposed Tech.
2001	13326	14112	12376	7.67 %	14.01 %
2002	14401	14817	13340	7.95 %	11.12 %
2003	14735	14680	14400	2.32 %	1.92 %
2004	16454	15444	14735	11.6 %	4.80 %
2005	17731	15908	16274	8.95 %	2.2 %
Average error %				7.71%	6.81 %

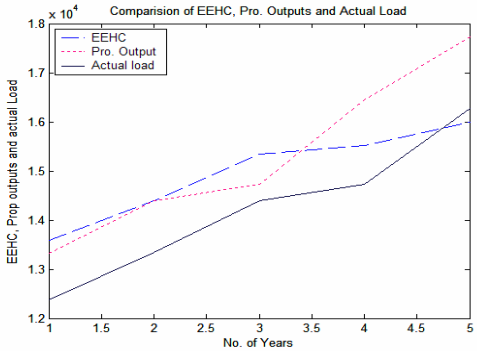

Fig.(3): Comparison Between EEHC, Forecasted Output and Actual Load of The Years (2001-2005)

This technique provides large percentage error regarding to annual year plan forecasting. Comparing between actual load, proposed ANN output and the output calculated by EEHC for five years are presented in Fig.(3). EEHC models and Proposed ANN produce average error that equal 7.71 % and 6.81 %, respectively, for consecutive five years. This result evidence the validity of ANN application within odd load curve than other methods.

CONCLUSIONS

From the previous illustrations the following conclusions can be deduced:

- Training of the network for all techniques implements the suggested goal performance but sometimes the output is not satisfied. The reasons of results failure can be summarized:
  1) Shortage of input parameters of the network structure, yields to inaccurate results especially in last years.
  2) Few number of observations of one parameter shares in difficulty of learning process.
  3) The untidy change in the load curve for year after year makes sure that load forecasting estimation requires more investigation to construct the effective network, and also for the other input parameters. Then, it must be investigated to find other factors that are responsible for odd shape of load curve
  4) Unavailable inputs may be predicted or guessed and this may yields to inaccurate outputs. So, accurate programs must be used to determine these values as the forecasted targets are based on them.
  5) Weather condition is one of the most important parameter affecting on medium-term load forecasting but it difficult work to gain.
- One year plan technique is preferable than five years plan to apply within few available inputs.
- Consumed time of training has very short period comparing with other estimating methods.
- In general, suggested goal is met and the performance has accurately achieved. Also, satisfied outputs of testing structure are gained except for some few ANNs structures that have inaccurate results.

REFERENCES

[1] Mohamed T. M. Youssef, 1999, "Load Forecasting Using ANN techniques.", Ph.D Thesis, Faculty of Engineering, Zagazig University.

[2] M.S. Kandil,S.M, El-Debeiky and N.E. Hasanien, , May, 2002, " Long-Term Load Forecasting for Fast Developing Utility Using a Knowledge-Based Expert System",IEEE Transaction on Power Systems, Vol.17, No. 2.

[3] S.A. Soliman, R.A. Alammari, M.E. El-Hawary and H.K. Temraz**.** April, 2004, "Long-Term Electric Peak Load Forecasting for Power System Planning: A Comparative Study", The Arabian Journal for Science and Engineering, Volume 29, Number1B.

[4] H. Demuth, and M. Beale, 1992, "Neural Network–Toolbox for Use with Matlab™", the Math Works, Inc.