

Load Forecasting in Thailand Using Neural Networks

Juntakan Taweekun

Department of Mechanical Engineering, Faculty of Engineering, Prince of Songkla University, Hatyai, Songkla, Thailand 90112

Abstract: In this study, gross domestic product, the gross national product, index of manufacturing production, number of population, number of households, oil price, fuel consumption, average ambient temperature and electricity consumption are used as inputs in neural networks to predict peak electric power demand. Three-layered feed-forward neural networks trained with a fast non-iterative training algorithm were used in this study for this long-term forecasting. Neural networks were optimized by removing redundant inputs. The results show that the peak electric power demand for 2003, 2005, 2010, 2015 and 2020 are 19,067 MW, 22,577 MW, 31,527 MW, 36,753 MW and 38,212 MW, respectively. Neural network can give the accurate estimations as indicated by the performance measures namely, root mean square error, coefficient of variation, mean bias error.

Keywords: neural network, forecasting, economic factor, electric power demand, weather climate, electricity consumption

1. INTRODUCTION

In Thailand, electricity consumption and peak electric power demand grew at a rate of 8.8% and 8.9% per annum during 1993 to 2003, respectively. Electric power demand will be continued increasing occurring in decreasing of the load factor of the power system. Figure 1 shows the electricity consumption and peak electric power demand in Thailand during 1993 to 2003. It was observed that the electricity consumption and peak electric power demand decreased sharply according to economic crisis during 1998 to 1999.

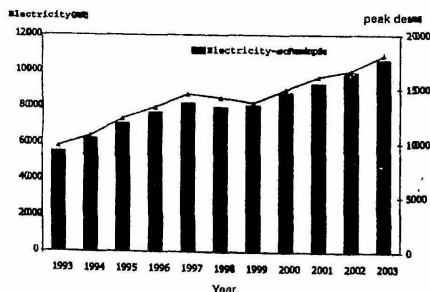


Figure 1: Electricity consumption and peak electric power demand in Thailand.

Load forecasting has become one of the most important aspects of electricity utility planning. Moreover, an accurate load forecasting is helpful in

developing a power supply strategy, financing planning, marketing research and electricity management. Most of the previous studies have been concerned with the application of different models based on different behavior assumptions about the load shape i.e. regression models^[1], expert system based models^[2], state space models, and time series models.

Alternative approach for load prediction is the use of neural networks^[3,4]. Nevertheless, most of the previous studies concentrated only on short-term load forecasting from a minute up to 24 hours. A few studies were focused on mid-term and long term load forecasting using neural network^[5,6]. Previous studies from mentioned references revealed that electric power demand depends on a number of factors that are non-linear characteristics and the appropriate approach for load prediction is the use of neural network.

This study focused on the long term prediction of peak electric power demand for Thailand. Neural network was used to train and test the developed model. Economic data, electricity consumption and weather climate data recorded for 10 years between 1993 and 2002 were used as inputs in neural network for model development and verification. Predictions were done for target years 2003, 2005, 2010, 2015 and 2020.

2. NEURAL NETWORK MODEL

Applicability of three-layered feed-forward neural network trained with modified version of Regulated Activation Weight Network (MRAWN) trained using an algorithm based on orthogonal least square (OLS) estimation were examined^[7]. The structure is shown in Fig.2. Hidden layer contains neurons with a

nonlinear activation function while the output neuron has a linear activation function.

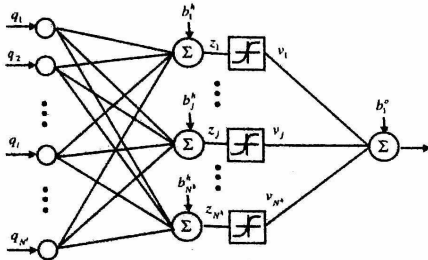


Figure 2: Three layered feed-forward neural network.

Each input to a neuron in the hidden layer is multiplied with an activation weight. These multiplied inputs are added together with a bias input. This is expressed by:

$$z_j(t) = \sum_{i=1}^{N^i} w_{ij}^h q_i(t) + b_j^h \quad (1)$$

The output of the hidden layer is then given by:

$$v_j(t) = a(z_j(t)) \quad (2)$$

The output of the complete neural network is obtained as:

$$o(t) = \sum_{j=1}^{N^h} w_j^o v_j(t) + b^o \quad (3)$$

where the inputs and outputs of the neural network are denoted by q and o . Subscript j denotes the j^{th} neuron in the hidden layer, and subscript i denotes the i^{th} input. Index t denotes the t^{th} event and the quantity N^i is the number of inputs to the network. The weight w_{ij}^h are called activation weights and the j^{th} activation bias is denoted as b_j^h . The function $a(\cdot)$ is the activation function which sigmoidal function is used in this study. The quantity N^h is number of hidden neurons. The weight w_j^o are the output weights and the output bias is denoted as b^o . The MRAWN training algorithm was applied to estimate the weights of neurons.

In the non-iterative training algorithm for Regulated Activation Weight Network, training is divided into two sub-problems: estimation of the activation weights, and estimation of output weights. The sub-problem of estimating activation weights is solved optimally by OLS estimation using a set of N_o ($N_o \gg N_h$) input output data pairs. Output weights are estimated using ordinary least squares. For estimation of the activation weights matrix, it is considered that a nonlinear function can be approximated by using a sequence of local linear functions. Training data are divided into subsets in an arbitrary way. The number of neurons is the same as the number of subsets of data.

The modified training algorithm proposed in [8] uses OLS fitting with input selection in place of the ordinary least square fitting. The automatic input selection algorithm selects the inputs one by one in the order of their relevancy until some termination criteria is met. This study follows the procedure described in [8] where classical Gram-Schmidt scheme is used for orthogonalization and a combination of Akaike and Bayesian information criteria is used as the terminating criteria. Outputs of the hidden neurons computed for all training data with already known hidden weights are then used to optimally estimate the vector of output weights through ordinary least square estimation.

3. PERFORMANCE MEASURES

This study uses three indicators to evaluate predicting peak electric power demand performances, namely the root mean square error (RMSE), the coefficient of variation (CV) and the mean bias error (MBE). These three indicators have been extensively used for estimation of the accuracy of neural networks and other black-box type models predicting electrical demand and energy consumption in buildings [3,9]. The RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{pred,i} - y_{actual,i})^2} \quad (4)$$

where $y_{pred,i}$ and $y_{actual,i}$ are the predicted and actual values of the dependent variable corresponding to the i^{th} set of inputs respectively, and n is the total number of observations. The CV and MBE are defined as:

$$CV = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_{pred,i} - y_{actual,i})^2}}{\bar{y}_{actual}} \quad (5)$$

$$MBE = \frac{\frac{1}{n} \sum_{i=1}^n (y_{pred,i} - y_{actual,i})}{\bar{y}_{actual}} \quad (6)$$

where $\bar{y}_{actual}$ is the mean of the actual values of the dependent variable.

4. INPUT SELECTION

As mentioned earlier, electric power demand depends on the economic factors and weather climate. In this study, the following factors are carefully selected for prediction of peak electric power demand.

1. Gross domestic product

GDP reveals the size of economic activity and economic conditions in the country.

2. Gross national product

GNP is used as a measure of the size of the nation's economy and as an indicator of economic trends.

3. *Index of manufacturing production*
Industrial movement can be expressed in term of index of manufacturing production.
4. *Number of population*
5. *Number of households*
Number of population and households affect directly in proportion to electric power demand.
6. *Oil price*
7. *Fuel consumption*
8. *Average ambient temperature*
At a peak load, a temperature rise of 1°C causes an increase in electric power of 4500 MW [6]. Thus, average ambient temperature as weather climate factor is also considered as input in neural networks.
9. *Electricity consumption*

The symbols used for each variable in Tables 2, 3 and 4 are shown in Table 1.

The values of independent and dependent variables during 1993 to 2002 for this study are given in Table 2. These data were obtained from each sector in Thailand namely, Bank of Thailand, Department of Industrial Works (Ministry of Industry), National Statistical Office (Ministry of Information and Communication Technology), the Meteorological Department, and Energy Policy and Planning Office (Ministry of Energy).

However some of the inputs could be irrelevant and should be removed looking at the prediction error

for test data. Thus the design of the neural network model should be done iteratively.

Table 1: Variables considered for predicting electric power demand.

Symbol	Independent variables
<i>GDP</i>	Gross domestic product
<i>GNP</i>	Gross national product
<i>INDEX</i>	Index of manufacturing production
<i>PO</i>	Number of population
<i>HO</i>	Number of households
<i>OIL</i>	Oil price
<i>FUEL</i>	Fuel consumption
<i>TEMP</i>	Average ambient temperature
<i>ELEC</i>	Electricity consumption
Dependent variables	
<i>PEAK</i>	Peak electric power demand

Initially, for training and testing in developing model, the data in Table 2 were normalized into the interval 0.05 to 0.95 by using the following formulation:

$$\text{Normalized data} = 0.9 \left(\frac{y - y_{\min}}{y_{\max} - y_{\min}} \right) + 0.05 \quad (7)$$

Where y is the actual values, $y_{\min}$ and $y_{\max}$ are the minimum and maximum of the variable, respectively.

Table 2: Values of independent and dependent variables.

Year	GDP ^a (MB)	GNP ^a (MB)	INDEX ^b	Number of Population ^c (person)	Number of Households ^c (Unit)	Oil Price ^c (B/litre)	Fuel Consum. ^c (Mlitre)	Avg. Temp. ^d (°C)	Electricity Consum. ^e (GWh)	Peak demand ^e (MW)
1993	617,727	609,181	86	58,336,072	13,336,167	8	2,517	27	4,603	9,730
1994	673,243	663,684	92	59,095,419	13,997,419	7	2,793	27	5,213	10,709
1995	735,434	725,188	100	59,460,382	14,697,085	8	3,140	27	5,906	12,268
1996	778,835	764,272	108	60,116,182	15,341,204	9	3,419	27	6,424	13,311
1997	768,154	752,079	107	60,816,227	15,495,755	9	3,405	28	6,840	14,506
1998	687,421	669,374	97	61,466,178	15,888,639	9	3,081	28	6,658	14,180
1999	717,997	703,877	109	61,661,701	16,248,890	9	3,076	27	6,733	13,712
2000	751,349	746,254	112	61,868,746	16,516,322	13	2,963	27	7,299	14,918
2001	765,926	759,634	114	62,308,887	16,910,473	13	2,902	27	7,739	16,126
2002	806,158	798,220	123	62,799,872	17,309,344	13	2,992	28	8,260	16,661

Source: ^aBank of Thailand, Thailand

^bDepartment of Industrial Works, Ministry of Industry, Thailand

^cNational Statistical Office, Ministry of Information and Communication Technology, Thailand

^dThe Meteorological Department, Thailand

^eEnergy Policy and Planning Office, Ministry of Energy, Thailand

MB-Million Baht (the currency of Thailand)

Base year of index of manufacturing production is year 1995 (equal 100)

Mlitre-Million Litre

GWh-Giga Watt Hour; MW-Mega Watt

4. DESIGN OF NEURAL NETWORKS FOR FORECASTING

Neural network was trained to predict peak electric power demand. The data in Table 2 are divided into two sets: training data and testing data as follows:

- Training data set: year 1993-1999 (7 yrs)
- Testing data set: year 2000-2002 (3 yrs)

The neural networks were trained using MRAWN training algorithm. Firstly, all independent variable in Table 1 were included in the input vector to the neural networks, some of them may be irrelevant for predicting peak electric power demand. Irrelevant inputs make neural network more complicate and increase the prediction error. In order to eliminate irrelevant inputs, effect of each variable was investigated.

In optimizing the input vector, for every input combination, a new neural network was required to be trained, and the number of hidden units of the network needed to be optimized. This requires training of large number of neural networks of

different configurations and was practicable due to use of fast non-iterative MRAWN training algorithm. Initially, for case 1, all independent variables were included and the results are given in Table 3. It was observed that electricity consumption and average ambient temperature are the most important factors for prediction as seen from the performance measures. Secondly, the effects of fuel consumption and oil price were investigated and found that the models are improved when these two factors are eliminated for case 4 and case 5, respectively.

Examination of the effects of number of population, the index of manufacturing production, gross national product and gross domestic product indicated that inclusion of these parameter gives best performance for peak electric power demand, as indicated by *RMSE*, *CV*, and *MBE* values. From these four parameter (number of population, the index of manufacturing production, gross national product and gross domestic product), it was observed that the most important parameter for predicting peak electric power demand is the number of population. The second and third are gross national product and gross domestic product, respectively. Index of manufacturing production is the last one.

Table 3: The performance of peak electric power demand models.

Case	Inputs	No. of hidden units	RMSE (MW)	CV (%)	MBE (%)
1	GDP, GNP, INDEX, PO, HO, OIL, FUEL, TEMP, ELEC	4	493	5.63	3.08
2	GDP, GNP, INDEX, PO, HO, OIL, FUEL, TEMP	9	1,084	10.34	0.48
3	GDP, GNP, INDEX, PO, HO, OIL, FUEL, ELEC	9	1,084	10.34	0.48
4	GDP, GNP, INDEX, PO, HO, OIL, TEMP, ELEC	3	431	3.71	2.24
5	GDP, GNP, INDEX, PO, HO, TEMP, ELEC	3	431	3.71	2.24
6	GDP, GNP, INDEX, PO, TEMP, ELEC	1	290	2.73	0.16
7	GDP, GNP, INDEX, TEMP, ELEC	6	632	4.86	1.35
8	GDP, GNP, PO, TEMP, ELEC	1	316	2.78	0.3
9	GDP, INDEX, PO, TEMP, ELEC	4	455	4.57	3.44
10	GNP, INDEX, PO, TEMP, ELEC	6	576	4.51	0.92

Finally, the best performances of prediction were obtained (case 6 in Table 3) when gross domestic product, gross national product, index of manufacturing production, number of population, average ambient temperature and electricity consumption are present. The corresponding *RMSE*, *CV* and *MBE* values for testing data are 290 MW, 2.73%, 0.16% for peak electric power demand. Figure 3 compares the results obtained from the predicting (neural network model) and actual data.

The values of average residual of peak electric power demand are approximately 151 MW or 0.93% discrepancy. The residual is the difference between peak electric power demand of actual data and predicted data (neural network model) as following:

$$\text{Residual} = Y_{\text{actual},i} - Y_{\text{pred},i} \quad (8)$$

where $Y_{\text{actual},i}$ and $Y_{\text{pred},i}$ are defined earlier.

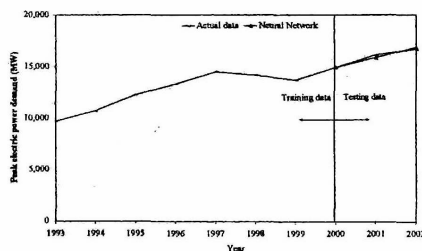


Figure 3: Predictions of peak electric power demand of 2000-2002 (training data: 1993-1999).

For load forecasting in 2005, 2010, 2015 and 2020, firstly, the model was tested and validated by using the developed model as mentioned earlier which compose of the gross domestic product, gross national product, index of manufacturing production, number of population, average ambient temperature and electricity consumption as inputs. Data from

1993 to 1999 was used for training to predict peak electric power demand in 2003 by using average incremental growth rate from the previous data. The comparison between actual data and predicted data using the incremental growth rate is given in Table 4. The discrepancy of the prediction is approximately 5.2% within acceptable ranges.

Table 4: Comparison between actual data and neural network of peak electric power demand in 2003.

Year	Inputs	Number of hidden units	Actual data (MW)	Neural network (MW)	Error (%)
2003	GDP, GNP, INDEX, PO, TEMP, ELEC	5	18,121	19,067	5.2

The developed model using incremental growth rate for predicting peak electric power demand in 2003 has good accuracy then this model will be used to forecast peak electric power demand in 2005, 2010, 2015 and 2020 with the same training data set and the same method. The results are given in Table 5 and Fig.5. These data are useful for applications i.e. generation capacity installation, investment capital cost, electricity price setting.

Table 5: Predictions of peak electric power demand.

Year	Actual data (MW)	Neural network (MW)
2005	-	22,577
2010	-	31,527
2015	-	36,753
2020	-	38,212

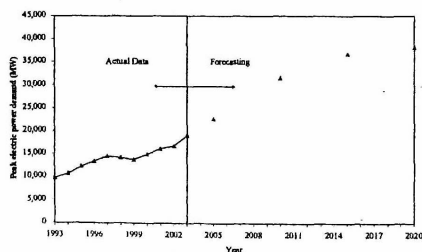


Figure 4: Predictions of peak electric power demand of 2005, 2010, 2015 and 2020. (Training data: 1993-1999).

5. CONCLUSION

Neural network was used to predict peak electric power demand and the results show that it can give the accurate estimations. Moreover, the electricity consumption and average ambient temperature mainly affect the prediction of peak electric power demand which indicated by the performance measures comparing with other, independent variables. The best performances of prediction were obtained when gross domestic product, gross national product, index of manufacturing production, number

of population, average ambient temperature and electricity consumption were included. Results of this study will be useful for other country's utilities.

REFERENCES

- [1] N. Hubele, C. Cheng, 1989. Identification of Seasonal Short-Term Load Forecasting Models Using Statistical Decision Functions. IEEE Power Engineering Society.
- [2] S. Rahman, A. Bhatnagar, 1970. An Expert System Based Algorithm for Short-Term Load Forecasting, Part 1: Forecasting Modeling. IEEE Trans Power Appar Syst, PAS. Vol 89. No 7. pp1678-82.
- [3] J. Taweekun, A. Rajapakse, S. Chirarattananon, 2002. A Study on Use of Neural Networks and DOE-2 Program for Predicting Cooling Load and Electrical Demand in Buildings. 1st International Conference on Sustainable Energy Technologies, Portugal. ppBE20/53-BE25/53.
- [4] D.C. Park, M.A. El-Sharkawi, R.J. Marks, L.E. Atlas, M.J. Damborg, 1991. Electric Load Forecasting Using an Artificial Neural Network. IEEE Trans Power Syst. Vol 6. No 2. pp442-449.
- [5] C.C. Hsu, C.Y. Chen, 2003. Regional Load Forecasting in Taiwan-Applications of Artificial Neural Networks. Energy Conversion and Management. Vol 44. pp1941-1949.
- [6] B. Kermanshahi, H. Iwamiya, 2002. Up to Year 2020 Load Forecasting Using Neural Nets, Electrical Power and Energy System. Vol 24. pp789-799.
- [7] A. Rajapakse, K. Furuta, S. Kondo, 1999. Identification of Nonlinear Dynamic Models with Partially Connected Neural Networks Trained Using Orthogonal Least Square Estimation. Trans. IEE of Japan. Vol 119. No 3.
- [8] H.A. Braake, H.J. Can, G. Straten, H.B. Verbruggen, 1995. Two Step Approach in Training of Regulated Activation Weights Neural Networks, Internal Report R95.043, Dept. of Electrical Engineering, Delft Univ. of Technology.
- [9] J.F. Kreider, & J. Haberl, 1994. Predicting Hourly Building Energy Usage: The Great Predictor Shootout-Overview and Discussion of Results, ASHRAE Transactions. Vol 100. No 2. pp1104-1116.