ANN BASED LONG-TERM SECTOR-WISE ELECTRICAL ENERGY FORECASTING

J. Kumaran Kumar[1] and G. Ravi[2]
[1]Department of Computer Science and Engineering, Pondicherry Engineering College, Puducherry, India
[2]Department of Electrical and Electronics Engineering, Pondicherry Engineering College, Puducherry, India
E-Mail: kumaran@pec.edu

ABSTRACT

Long term load forecasting is an important aspect of electric utility resource planning and utility expansion. This paper presents an ANN based forecasting model that predicts the sector-wise electrical energy demand in India for the future years up to 2025. The model requires per capita GDP and population and offers the sector-wise electrical energy demand forecast. The comparisons of the results with that of RM exhibit the effectiveness of the proposed model.

Keywords: load forecasting, artificial neural networks.

1. INTRODUCTION

The electric utilities as well the government is responsible for providing adequate power supply to the consumers and finding difficult to supply the electrical energy demand, which exponentially increases to meet the development of high technology industries and high usage of domestic appliances due to economic growth, in developing countries like India. In contract, the existing thermal power plants that pollute the atmosphere require to be closed and cleaner power technologies are to be pursued to meet the requirements of environmental regulations and the demand for clean air by the public. Under these circumstances, the functioning of Koodamkulam nuclear power project in India was delayed due to the opposition caused by the public, resulting in a huge revenue loss. It is therefore very important to the government as well as to the utilities to have advance knowledge of the future electrical energy demand, so that the load can be met without any interruption.

Load forecasting (LF) is an integral part in efficient planning, operation and maintenance of a power system. Load forecast can be classified into, depending on the area of application, long-term (1-10 years), medium-term (1 day to 1 year) and short-term (up to one day) [1, 2]. Each term of forecast has its own merits to the utilities. Long term forecasting is essential for system planning with a view of inflating the system capability in order to meet the long term growth in demand. Medium term LF is necessary for the scheduling of fuel supply and maintenance operations and planning for inter utility power transfer. Short-term LF is necessary in the daily operation such as unit commitment, energy transfer scheduling, fuel scheduling and demand side management. However, neither the accurate forecast nor the production of electrical energy is as easy as it looks, because (a) forecasting may be inaccurate (b) peak demand depends on temperature (c) data required for forecasting such as weather, economic data are not available and (d) construction of new power plants and transmission facilities require huge investment and take several years for completion.

The load demand is estimated by extrapolating the relationship between load and variables such as temperature, time, etc that are strongly tied. Determining the relationship is a two stage process that requires: a) identifying the relationship between the load and the related variables and b) quantifying the relationship through the use of a suitable parameter estimation technique.

Many techniques such as auto regressive integrated moving average [3] and regression analysis (RA) [4-7] have been investigated to solve the problem of LF in the last few decades. Recently, considerable interest appears to be focused on the application of artificial neural networks (ANN) for LF due to their ability to extract the relationship among input variables and output through learning from the available database [8-12]. ANNs combined with RA [13-16] as well with fuzzy logic [17, 18] for LF have been outlined. The other hybrid versions for LF have also been notified in [19, 21]. Most of the studies focus on short-term LF. Only a few studies have been carried out for medium and long term LF.

An ANN based model involving Per capita GDP and Population for forecasting India's sector-wise electrical energy demand for future years unlike the existing models predicting the net energy demand is suggested. The paper is organized as follows: section II outlines the existing regression model (RM), section III explains the proposed model (PM), section IV provides the simulation results and section V concludes.

2. REGRESSION MODEL

Traditional models for LF can be generally classified as time series models and RMs. In time series models, the previous load is extrapolated to obtain future loads. These models are frequently augmented with transfer functions in order to adapt the consumer response to changing weather patterns and intangible factors. These models require large amount of data and a complex estimation procedure. RMs constitute the second major modeling technique, wherein the database is frequently divided into smaller segments, whereby a RM is built for each segment, such as a season or a day or a week.

VOL. 10, NO. 1, JANUARY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

RA is a technique used for analyzing the numerical data. The dependent variable y_i is a linear combination of the parameters, α and the independent variables, x_i which could be linear or nonlinear. The simple linear and multiple linear regressions are the two basic types of linear regression. For instance, in simple regression of N data points' modeling, there is one independent variable, x_i, and two parameters, α_0 and α_1, which yield a straight line, called fitted regression line:

$$y_i = \alpha_0 + \alpha_1 x_i + e_i \qquad i = 1, \cdots, N \tag{1}$$

In multiple linear regressions, there are more than one independent variables or function of independent variables. For example, the preceding regression with x_i^2 term gives a parabola:

$$y_i = \alpha_0 + \alpha_1 x_i + \alpha_2 x_i^2 + e_i \qquad i = 1, \cdots, N \tag{2}$$

Although the right-hand side expression is quadratic, it is still considered to be linear regression, as it involves linear parameters, α_0 α_1 and α_2 .

In the general multiple RMs, there may be m independent variables:

$$y_i = \alpha_0 + \alpha_1 x_{1i} + \cdots + \alpha_m x_{mi}^2 + e_i \qquad i = 1, \cdots, N \tag{3}$$

Where e_i is the error term, which represents the unexplained variation in the dependent variable and is treated as a random variable?

In practice, the performance of RM depends on the form of the data-generating process and its relation to the regression approach used. Typically, the best fit is evaluated by using the least squares method, although other criteria are also used.

3. PROPOSED MODEL

The goal of the PM is to forecast the sector-wise electrical energy demand in future years with minimum input data. Recently ANNs find extensive acceptance in many disciplines for modeling complex real-world problems that includes LF because of their clear and easy model implementation artifact. They are like human brains and massively parallel-distributed information processing systems with highly flexible configuration possessing excellent nonlinearity capturing ability. Usually, they are multi-layer feed forward networks comprising of an input layer, an output layer and a hidden layer, each with a set of neurons and designed to perform a particular task. Solving undefined relationships between input and output variables, approximating complex nonlinear functions and implementing multiple training algorithms are the better-known advantages of these tools [8, 9]. The proposed forecasting model is thus based on ANN.

There are a large number of input data such as such as weather, average temperature, time, number of households, number of air conditioners, amount of CO2 pollution, oil price, economy, population, etc., which are related to the electrical energy demand in any country. It would be difficult to train a neural network with all the available input data, as the number of connection weights and neurons would be extremely large. It is thus essential to reduce the number of inputs to a neural network and select an optimum number of mutually independent inputs, which are able to clearly establish the input-output relationship. Besides, many of these factors are usually used only for short term forecasting of energy demand.

Among these factors, the population growth as well the continuous improvement in the public revenue and living standards, represented through per capita GDP, are linked with the total energy consumption of any country [11, 13]. The plots relating the energy consumed with per capita GDP and the population in India during the years 1980-2012, shown through Figures 1 and 2, clearly indicate that the energy consumed in India increases with both the per capita GDP and the population.

In the light of the fact that the per capita GDP and population establish a good relationship with the electrical energy demand, they are considered as inputs in the PM. The structure of PM is shown in Figure-3. The model requires two inputs and offers six outputs ($E_1 - E_6$), representing energy demand in the sectors of industrial, agricultural, domestic, commercial, railways and other areas. The historical data set comprising the input (X) and the target (T) vectors for the PM are as follows.

$$\{X \leftrightarrow T\} = \{Pop, \mathrm{GDP} \leftrightarrow E_1, E_2, E_3, E_4, E_5, E_6\} \tag{4}$$

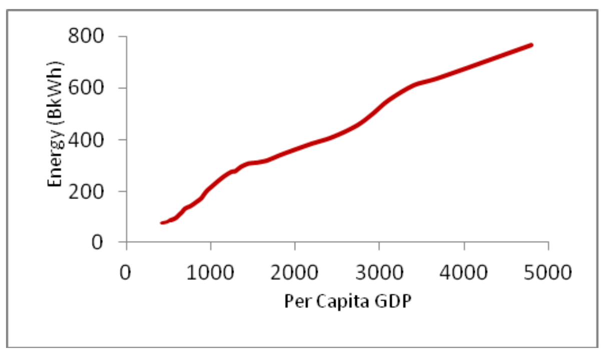

Figure-1. Per capita GDP versus energy.

VOL. 10, NO. 1, JANUARY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

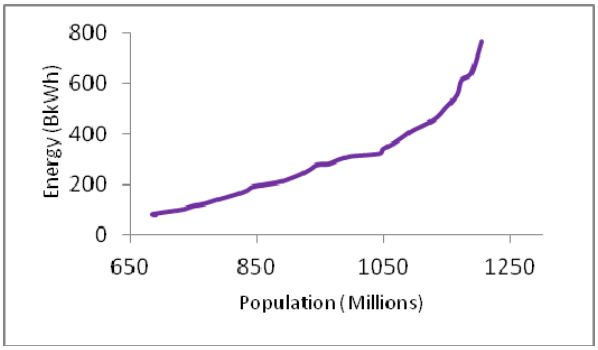

Figure-2. Population versus energy.

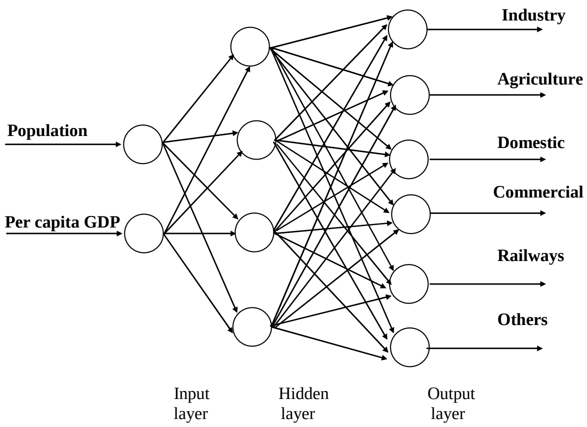

Figure-3. Proposed ANN based forecasting model.

The generated input-target data are split into two partitions: the first one is the training data, which is used to train the network and the second, the testing data, is used to assess how well the network is generalized. There is a possibility of obtaining good performance on the training data followed by much poor performance on the test data. This can be avoided by ensuring that the training data is uniformly distributed.

During training of the neural network, higher valued input variables may tend to suppress the influence of smaller ones. Besides, if the raw data is directly applied to the network, there is a risk of the simulated neurons reaching the saturated states. If the neuron becomes saturated, then the changes in the input value will produce a very small change or no change in the output value. This affects the network training to a great extent. The raw data is therefore normalized before it is applied to the neural network. One way to normalize the data x is by using the expression:

$$x_n = \frac{(x - x_{\min}) \times (U_R - L_R)}{x_{\max} - x_{\min}} + L_R \qquad (5)$$

Where x_n is the normalized value?

$x_{\min}$ and $x_{\max}$ are the minimum and maximum values of the variable x respectively

L_R and U_R lower and upper range for normalization respectively

The developed ANN model is trained by back-propagation. Tangent hyperbolic and linear functions are chosen as the activation function for the hidden layer neurons and the output neurons respectively. During training, the connection weights are adjusted to correctly map the training set vectors at least to within a defined error limit. A trial and error procedure is adopted in selecting the number of hidden layers and hidden neurons in such a way to obtain the ANN model that correctly predicts the energy demand for the testing data. Once the network is trained, the weights are frozen and the network is ready to forecast the sector-wise electrical energy demand for any unforeseen input data.

4. SIMULATIONS

Successful operation of ANN based load forecasters requires an appropriate training data set that can adequately cover the entire solution space with a view to recognize and generalize the relations among the problem variables. In this work, a historical data during the period of 1980-2012 have been used. The actual India's sector-wise energy consumption, the per capita GDP and the population data have been taken from the references [22-24]. The per capita GDP and population, calculated through RA during the years 2013-2025 and presented in Table-2, are considered as input data for actuating the PM, after training. The goodness of the forecast is evaluated through the following mean absolute percent error (MAPE).

$$\text{MAPE} = \frac{\sum_{i=1}^{nfo} |\Phi_i|}{no} \qquad (6)$$

Where no is the number of forecasted outputs?

$$\Phi_i = \frac{Actual\ \ Energy_i - Forecasted\ \ Energy_i}{Actual\ \ Energy_i} \times 100$$

VOL. 10, NO. 1, JANUARY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

Table-1. Comparison of results for yesteryears.

Year	Per capita GDP	Population (Millions)		Forecasted electrical energy (BkWh)						MAPE	
				Industry	Agriculture	Domestic	Commercial	Railways	Others	RM	PM
1989	821.48	817.49	Actual	76.82	38.85	24.61	10.06	4.04	5.82		
			RM	82.66	43.15	24.70	10.31	3.94	6.58	6.18	
			ANNBP	76.79	42.18	26.32	10.24	3.97	6.93		6.36
1997	1285.94	962.38	Actual	104.17	84.02	55.27	17.52	6.53	12.64		
			RM	97.40	80.33	56.46	16.63	6.58	12.74	3.28	
			ANNBP	99.71	82.62	55.63	16.42	6.58	13.33		3.18
2005	2190.27	1080.26	Actual	137.59	88.56	95.66	31.38	9.49	23.45		
			RM	148.18	95.38	97.21	34.09	9.56	24.23	4.95	
			ANNBP	137.40	93.29	95.69	33.29	9.32	23.17		2.43
2009	3103.73	1166.08	Actual	209.47	109.61	131.72	54.19	11.43	37.58		
			RM	222.44	115.08	137.31	53.87	12.10	35.31	4.65	
			ANNBP	206.95	110.72	132.98	55.89	12.47	33.65		4.31
Average MAPE										4.77	4.07

Table-2. Results of the RM.

Year	Per capita GDP	Population (Millions)
2013	5125.28	1209.75
2014	5690.23	1226.25
2015	6295.78	1248.56
2016	6938.99	1267.88
2017	7615.92	1287.75
2018	8321.45	1309.94
2019	9049.44	1330.69
2020	9792.53	1351.38
2021	10542.00	1372.50
2022	11287.82	1390.13
2023	12018.59	1406.69
2024	12721.35	1418.81
2025	13381.58	1431.25

ARPN Journal of Engineering and Applied Sciences

Table-3. Results of RM.

Year	**Forecasted electrical energy** (BkWh)					
	Industry	**Agriculture**	**Domestic**	**Commercial**	**Railways**	**Others**
2013	316.59	149.30	201.72	85.90	16.21	47.50
2014	348.98	163.18	225.37	96.65	17.29	52.67
2015	380.66	179.39	247.70	105.76	18.84	57.92
2016	417.05	197.25	279.28	117.44	20.39	63.88
2017	452.81	216.18	306.93	126.56	21.81	69.52
2018	489.81	234.45	344.48	140.50	23.33	75.95
2019	526.79	260.81	383.65	152.72	25.00	81.34
2020	558.12	287.67	422.26	164.58	26.89	89.44
2021	597.80	316.08	460.18	176.29	28.81	95.05
2022	621.94	347.88	502.34	191.56	30.67	101.78
2023	651.13	379.32	545.97	205.76	32.34	106.04
2024	672.23	417.67	594.35	218.98	34.41	111.39
2025	685.12	454.96	635.92	233.11	35.88	114.89

One of the challenges in the design of ANN is the proper selection of the number of neurons in the hidden layer, which affects the learning capability and leads to the complexity of the problem. The fundamental rule is to select the minimum number of hidden neurons just enough to ensure the complexity of the problem, but too many may cause over fitting of the training set and losing the generalization ability. In this paper, a trial and error scheme has been adopted in determining the appropriate number of hidden neurons. The number of neurons in the hidden layer is varied from 2 to 10 and the resulting MAPEs are compared for the testing data. It is found that a hidden layer with 7 neurons gives satisfactory results. The results of the PM are compared with that of RM with a view to evaluate the performance of the PM.

The results for the chosen yesteryears are given in Table-1. It clear from the results that the MAPE of the PM is lower than that of RM. The lowered MAPE validates the model and ensures the accuracy of the results for future years. The results of RA and PM during the years 2011-2025 are presented in Tables 3 and 4 respectively. It is seen from these Tables that the sector wise energy demands, given by PM, are in general lower than that of the RM. This is graphically illustrated by comparing the net energy demand of all the methods over the forecasting period in Figure-4. The PM predicts the energy demand in the future years to be lower than that of RM and indicates the policy makers to allocate little lower funds for construction of new power plants and transmission systems with a view to meet the future energy demand.

Table-4. Results of PM.

Year	**Forecasted electrical energy** (BkWh)					
	Industry	**Agriculture**	**Domestic**	**Commercial**	**Railways**	**Others**
2013	320.14	150.47	201.05	84.60	16.05	47.13
2014	352.31	164.51	222.73	93.81	17.38	51.45
2015	385.66	180.51	246.88	103.61	18.85	56.10
2016	419.92	198.55	273.25	113.96	20.46	61.07
2017	454.55	218.69	301.94	124.85	22.20	66.36
2018	488.98	240.93	333.20	136.25	24.03	71.90
2019	522.73	265.29	366.66	148.19	25.96	77.70
2020	555.16	291.74	402.34	160.61	27.91	83.63
2021	585.53	320.24	440.03	173.43	29.87	89.61
2022	613.48	350.68	479.38	186.71	31.77	95.55
2023	638.14	382.95	520.20	200.50	33.56	101.29
2024	658.84	416.90	562.05	214.66	35.11	106.70
2025	675.13	452.36	604.32	229.36	36.35	111.58

VOL. 10, NO. 1, JANUARY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

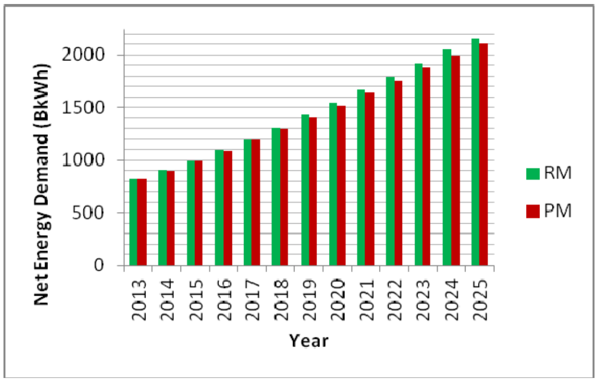

Figure-4. Comparison of net energy demand.

5. CONCLUSIONS

LF has been an important event in efficient planning and control of electric utilities ever since the conception of forecasting. Load serving entities use load forecasts for ensuring system security, scheduling generator maintenance, making long-term investments in generation and arriving at the most cost-effective merit order dispatch. The sector-wise electrical energy demand of India has been forecasted for the future years through the population and the socio-economic factors of per capita GDP by the developed ANN model trained through back propagation algorithm. It has been found from the results that the PM offers more accurate predictions than that of RM, helps the policy makers for allocating appropriate funds for constructing new generation plants and transmission systems to meet the future demands and attempts to offer reliable service to the customers in the future years.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the authorities of SCSVMV and Pondicherry Engineering College for their continued support, encouragement and the facilities provided to carry out this work.

REFERENCES

[1] P.E. McSharry, S. Bouwman and G. Bloemhof. 2005. Probabilistic forecasts of the magnitude and timing of peak electricity demand. IEEE Transactions Power Systems. 20: 1166-1172.

[2] Willis. H.L and Northcote-Green. J.E.D. 1984. Comparison tests of fourteen distribution load forecasting methods. IEEE Trans. on Power Apparatus and Systems. 103(6): 1190-1197.

[3] S. Gonzales Chavez, J. Xiberta Bernal and H. Llaneza Coalla. 1999. Forecasting of energy production and consumption in Asturias (Northern Spain), Energy. 24(3): 183-198.

[4] Arash Ghanabri, S.Farid Ghaderi and M.Ali Azadeh. 2010. Adaptive neurofuzzy inference system vs. regression based approaches for annual electric load forecasting. 2nd International Conference on Computer and Automation Engineering (ICCAE), Singapore, 26-28 February. 5: 26-30.

[5] C.L. Hor, SJ. Watson and S. Majithia. 2005. Analyzing the impact of weather variables on monthly electricity demands. IEEE Transactions on Power Systems. 20: 2078-2085.

[6] D. Asber, S. Lefebvre, J. Asber, M. Saad and C. Desbiens. 2007. Nonparametric short-term load forecasting. International Journal of Electrical Power and Energy Systems. 22(8): 630-635.

[7] Geoffrey K.F. Tso and Kelvin K.W. Yau. 2007. Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks. Energy. 32(9): 1761-1768.

[8] Henrique Steinherz Hippert, Carlos Eduardo Pedreira, and Reinaldo Castro Souza. 2001. Neural networks for short-term load forecasting: a review and evaluation. IEEE Transactions on Power Systems. 16(1): 44-55.

[9] Kermanshahi B and Hiroshi Iwamiya. 2002. Up to year 2020 load forecasting using neural networks. Electrical Power and Energy Syst. 24(9): 789-797.

[10] Madasu hanmandlu and Bhavesh Kumar Chauhan. 2011. Load forecasting using hybrid models. IEEE Trans. on Power Systems. 26(1): 20-29.

[11] S. Saravanan, S. Kannan and C. Thangaraj. 2012. Forecasting India's electricity demand using artificial neural network. IEEE- International Conference on Advances in Engineering, Science and Management (ICAESM -2012). pp. 79-83.

[12] Santosh Kulkarni, Sishaj P. Simon and K. Sundareswaran. 2013. A spiking neural network (SNN) forecast engine for short-term electrical load forecasting. Applied Soft Computing. 13: 3628-3635.

[13] A. Ghanhari, A. Naghavi, S.F. Ghaderi and M. Sabaghian. 2009. Artificial neural networks and regression approaches comparison for forecasting Iran's annual electricity load. Proceeding of International Conference on Power Engineering Energy and Electrical Drives. pp. 675-679.

[14] D.B.L. Bong, J.Y.B. Tan and K.C. Lai. 2008. Application of multilayer perceptron with back propagation algorithm and regression analysis for long-term forecast of electricity demand: a comparison. ICED International Conference on Electronic Design. pp. 1-5.

[15] Adem Akpinar. 2011. Modeling and forecasting of Turkey's energy consumption using socio-economic and demographic variables. Applied Energy. 88(5): 1927-1939.

[16] V. Sackdara, S. Premrudeepreechacharn and Naamsanroaj K. 2010. Electricity demand forecasting of Electricite DU Lao (EDL) using neural networks, TENCON-2010. pp. 640-644.

[17] Toly Chen. 2012. A collaborative fuzzy-neural approach for long-term load forecasting in Taiwan, Computers and Industrial Engineering. 63(3): 663-670.

[18] Toly Chen and Yu-Cheng Wang. 2012. Long-term load forecasting by a collaborative fuzzy-neural approach. Electrical Power and Energy Systems. 43(1): 454-464.

[19] Jie Wu, Jianzhou Wang, Haiyan Lu, Yao Dong and Xiaoxiao Lu. 2013. Short term load forecasting technique based on the seasonal exponential adjustment method and the regression model. Energy Conversion and Management. 70: 1-9.

[20] Wei-Chiang Hong, Yucheng Dong, Wen Yu Zhang, Li-Yueh Chen and B. K. Panigrahi. 2013. Cyclic electric load forecasting by seasonal SVR with chaotic genetic algorithm. Electrical Power and Energy Systems. 44: 604-614.

[21] Cheng-Ting Lin, Li-Der Chou, Yi-Ming Chen and Li-Ming Tseng. 2014. A hybrid economic indices based short-term load forecasting system, Electrical Power and Energy Systems. 54: 293-305.

[22] Data for per capita GDP available at http://www.indexmundi.com/india/gdp_per_capita_(ppt).html.

[23] Data for population available at http://www.populstat.info/Asia/indiac.htm.

[24] India’s sector wise energy demand available at http://www.iasri.res.in/agridata/08data%5Cchapter3%5Cdb2008tb3_57.pdf.