General review of long-term electric load forecasting models

Prof. F. González-Longatt, MIEEE
Research Associate
The University of Manchester
School of Electrical & Electronic Engineering
Ferranti Building A3, Sackville Street
Manchester M60 1QD
United Kingdom
Email: francisco.gonzalez-longatt@manchester.ac.uk
15th February of 2009

Introductions

Load forecasting is a central and integral process in the planning and operation of electric utilities. Long-term load forecasting helps an electric utility to make important decisions in a long-term horizon. In this paper a brief description of parametric and artificial intelligence based methods most used in long-term load forecast are shown. Mainly parametric methods: trend analysis, end-use modeling and econometric modeling are briefly discussed and compared. A general overview of main artificial intelligence based methods is shown.

Key words: Forecast Long-term, demand forecasting, Neural networks, Genetic Algorithms, Fuzzy Rules, Wavelet networks.

INTRODUCTION

Accurate long-term electric load forecasting plays an essential role for electric power system planning. It corresponds to load forecasting with lead times long enough to plan for long-term maintenance, construction scheduling for developing new generation facilities, purchasing of generating units, developing transmission and distribution [1].

The time horizon for long-term forecasting ranges between a few months and several years. Unfortunately, it is difficult to forecast load demand accurately over a planning period of this length [1]. This fact is due to the uncertain nature of the forecasting process. There are a large number of influential factors that characterize and directly or indirectly affect the underlying forecasting process; all of them are uncertain and uncontrollable. Therefore, any log-term forecast, by nature, is inaccurate. Most of the electric-load forecasting methods are dedicated to short-term (a few minutes to 24 h) forecasting but not as much for the long-term (1–10 years) or intermediate-term (a few days to several months) load forecasting.

Accurate long-term demand forecasting plays an essential role for electric power system planning. In this paper a general review of the main long-term electric load forecast models is developed. A general discussion over parametric and artificial intelligence based methods is shown and a general comparative table is developed to parametric models. Finally some important aspects about these models are gathered in the conclusion section.

Long-term electric Load forecasting methods classification

Generally, long-term load demand forecasting methods can be classified in to two broad categories:

1. *Parametric methods* and
2. *Artificial intelligence* based methods.

The *parametric methods* are based on relating load demand to its affecting factors by a mathematical model. The model parameters are estimated using statistical techniques on historical data of load and it's affecting factors. Parametric load forecasting methods can be generally categorized under two approaches:

1.1 Regression methods, and

1.2. Time series prediction methods.

The artificial intelligence based methods are further classified in:

2.1. Neural networks methods,

2.2. Generic algorithms methods,

2.3. Wavelet networks methods, and

2.4. Fuzzy logics methods.

1. Parametric Methods

The three types of well-known parametric methods are as, *trend analysis*, *end-use modeling* and *econometric modeling*.

1.1. Trend analysis models

Trend analysis extends past rates of electricity demand in to the future, using techniques that range from hand-drawn straight lines to complex computer-produced curves. These extensions constitute the forecast. Trend analysis focuses on past changes or movements in electricity demand and uses them to predict future changes in electricity demand [1].

1.2. End-use models

The end-use approach directly estimates energy consumption by using extensive information on end users, such as applications, the customer use, their age, sizes of houses, and so on. Statistical information about customers along with dynamics of change is the basis for the forecast [1].

End-use models focus on the various uses of electricity in the residential, commercial, and industrial sector. These models are based on the principle that electricity demand is derived from customer's demand for light, cooling, heating, refrigeration, etc. Thus, end-use models explain energy demand as a function of the number of applications in the market [1].

The basic structure of a definitional equation that is applicable to all end-uses has a simple form given by [7]:

$$D_{ij} = N_i C_i F_{ij} \tag{1}$$

where D_{ij} is the demand at hour j by end-use component i, N_i is the number of use components of type i, C_i is the connected load per use component of type i, and F_{ij} is the fraction of the connected load of use component i which is operating at hour j. The energy consumption over any period is calculated by summing over the corresponding demand values. Thus:

$$E_i = \sum_{j=1}^{m} D_{ij} \tag{2}$$

where E_i is the energy consumptions by end-use component *i*, and D_{ij} is demand at hour *j* by the end-use component *i*. The above basic forecasting equation merely states that the total demand by a given end-use is the product of the total connected load and the fraction of the connected load that is operating at the given time. The total connected load in turn is the product of the number of end-uses and the corresponding average connected load per end-use. Repeating these calculations for each hour of the day and for all end-uses yields the daily load profile for the system.

The above equations are applicable to residential, commercial and industrial sector. The equations require sane modification for the industrial and commercial sectors.

The end-uses of electricity in the residential are the major household appliances. Thus the basic forecasting equation converts to [7]:

$$R_{ij} = NA_i CA_i FA_{ij} \tag{3}$$

Where RD_{ij} is the demand by appliance of type *i* during hour *j*, NA_i is the number of appliances of type *i*, CA_i is the connected load of appliances of type *i*, and FA_{ij} is the fraction of appliances of type *i* operating during hour *j*. The total residential demand is merely the sum of the corresponding demand projections for the individual appliance types. The number of appliances is a function of the number of households and the appliance saturation. In some cases is convenient split the type of dwelling in: urban, sub-urban and rural population.

The fraction of the connected load of an appliance that is operating at a specific hour is the use factor of the appliance for that hour. The set of the use factors for a given day type and month describes the expected operating pattern of the appliance for that day/ The collection of use factors necessary to describe the operation of an appliance throughout the year is *the use matrix of the appliance* [7].

1.3. Econometric models

The econometric approach combines economic theory and statistical techniques for forecasting electricity demand. The approach estimates the relationship between energy consumption (dependent variables) and factors influencing consumption [1].

These models relate the demand for electric energy to economic and demographic variables such as population, income, and electricity rates. In this paper, the authors develop an error analysis [2].

The relationships are estimated by the *least square method* or *time series methods*.

One of the options in this framework is to aggregate the econometric approach, when consumption in different sectors (residential, commercial, industrial ,etc.) is calculated as a function of weather, economic and other variables, and then estimates are assembled using recent historical data. Integration of the econometric approach in to the end-use approach introduces behavioral components in to the end-use equations [1].

Econometric load forecasts of demand are based on a model whose parameters are estimated from historical data. A *linear* or *log-linear model* is usually postulated.

In the linear model, the variable to be forecasted, the endogenous variable, is a linear function of the exogenous variables; the exogenous variables being the "input" variables. In general, a linear dependence of the endogenous variable upon past (lagged) values of itself and of the exogenous variables may also be accommodated by the linear model.

In the log-linear model, a linear relationship is assumed between the logarithm of the endogenous variable and logarithms of the exogenous variables.

Log-linear models are the ones most commonly used for modeling the demand for electric energy.

Either the linear or log-linear models may be expressed by the following equation [2]:

$$y_i = \sum_{j=1}^{k} a_j y_{i-j} + \sum_{j=1}^{m} b_j u_{j,i} + c + v_i \tag{4}$$

where: y_i is either the annual energy sales or the logarithm of annual energy sales for year i. $u_{j,i}$ is the j-th exogenous variable, or logarithm of the *j-th* variable for year i. $v(t_i)$ is a random deviation for year i. a, b, c are model coefficients.

The model coefficients are estimated using historical data.

The construction of econometric models is usually accomplished on a trial and error basis. Exogenous variables, which would logically influence demand, are selected, a linear regression is performed, and statistical tests are made. Among the tests ordinarily performed are the *F-test* on the ratio of the variance explained by the regression to the unexplained variance, the *t-test* on the model coefficients to determine their significance, and the *Durbin-Watson test* to test for serial correlation of the residual errors. [3]

Frequently there may be more than one acceptable model. In that case the final choice may be made based on the plausibility of the exogenous variables as determinates of demand, and on the availability of forecasted values of the exogenous variables. It is the position of the authors, that a statistical analysis of forecasting errors should also be included in this modeling procedure [2].

The reason for this is that it is entirely possible to develop econometric models which fit the historical data well but which are very sensitive to errors in the model coefficients and to errors in the forecasted exogenous variables. Such models lead to unacceptably large, errors in forecasted demand [2].

Table 1. Comparative matrix between end-use and econometric models for medium and long term load forecast

Model	*Advantage*	*Drawback*
Trend analysis	• Simple. • Quick to implementation and solution. • Inexpensive to perform.	• A trend forecast is that it produces only one result, future electricity demand.
End-Use	• Very accurate. • Require less historical data than econometrics	• Sensitive to the amount and quality of end-use data. • Require more information about consumer and their equipments. • Problems to conduct next-year forecasts for sub-areas.
Econometric	• It provides detailed information on future levels of electricity demand, why future electricity demand increases, and how electricity demand is affected by all the various factors. • Combines economic theory and statistical techniques. • Simple methods could be used: least-squares method or time series methods.	• Require great amount of historical data. • Problems to conduct next-year forecasts for sub-areas. • Not exact for because only a finite number of data points, and thus because of the random component. • Model is not structurally perfect. There are additional exogenous variables which have influence on demand but are not included in the model. • In order to be accurate, the changes in electricity remain the same in the forecast period as in the past.

2. Intelligence Artificial based Methods

2.1. Artificial neural networks

Artificial neural networks (ANNs) have succeeded in several power system problems. The ANNs ability in mapping complex non-linear relationships is responsible for the growing number of its application to load forecasting [1], [8], [9]. Most of the ANNs have been applied to short-time load forecasting. Only a few studies are carries out for long-term load demand forecasting [10].

The design of neural network architecture involves decision making on type, size and number of neural being used [10]. The result of Output ANN is in (2).

$$y_i = \sum_{j=1}^{n} w_j x_j \tag{5}$$

Where $i = 1, 2, \dots, n$. x_j is input w_j is weight of network and y_i is one of the ANN's outputs.

Two important aspects in ANN are choosing the network architecture and method of training. There are two types which can be useful for long-term load demand forecasting: (i) *Recurrent neural network* (RNN) for forecasting the peak load, (ii) *feed-forward back propagation* (FFBP) for forecasting the annual peak load.

2.2. Wavelet networks

Wavelet theory is introduced to power load forecasting recently and received wide attention [1]. Comparing to traditional load forecasting methods, wavelet theory provides powerful and flexible tool to decompose load data into different frequency components, making it possible to analyze the characteristics of each component and improve forecasting accuracy. Wavelet packet analysis is the extension of wavelet analysis and it has better frequency resolution [4].

2.3. Genetic Algorithms

Genetic Algorithms (GAs) have recently received much attention as robust stochastic search algorithms for various problems [1]. This class of methods is based on the mechanism of natural selection and natural genetics, which combines the notion of survival of the fittest, random and yet structured, search and parallel evaluation of the points in the search space [5].

Genetic algorithms are a numerical optimization technique. They combine a Darwinian survival-of-the-fittest strategy with a random, yet structured information exchange among a population of artificial "chromosomes" [1]. This technique has gained popularity in recent years as a robust optimization tool for a variety of problems in engineering, science, economics, finance, etc. Some of the attractive features can be summarized as: Learning, Generic Code Structure, Optimality of the Solutions, Advanced Operators [5].

2.4. Fuzzy Logic Model

Fuzzy control systems are rule-based systems in which a set of so-called fizzy rules represents a control decision mechanism to adjust the effects of certain stimulus. The aim of fuzzy control systems is normally to replace a skilled human operator with a fuzzy rule-based system [1]. The fuzzy logic model provides an algorithm, which can convert the linguistic strategy based on expert knowledge into an automatic strategy. The fuzzy logic method is applied for scoring. The application of fuzzy rules will improve the model accuracy by avoiding arbitrariness for the purpose of the stud [6]. One of the applications of the fuzzy rules is to combine them with neural network to train ANN and have a better load demand forecasting. The benefit of the proposed hybrid structure was to utilize the advantages of both, i.e., the generalization capability of ANN and the ability of fuzzy inference for handling and formalizing the experience and knowledge of the forecasters. It has been demonstrated that the method give relatively accurate load forecasts for the actual data [6].

Conclusions

Accurate long-term electric load forecasting plays an essential role for the economic optimization and secure operation of electric power systems.

The confidence levels associated with classical forecasting techniques, when applied to forecasting problem in mature and stable utilities are unlikely to be similar to those of dynamic and fast growing utilities. All of these methods can forecast the load of the power system, but the amount of previous data and such variables which they need to forecast, make them different in accuracy from area to area.

Traditional methods such as trend, econometric and end-use methods exhibit reasonably reliable results. Models based on artificial intelligence can solve nonlinear problems, and because of nonlinear behavior of load, so they can be useful for long-term load forecasting.

References

[1] J.L. Ghods, M. Kalantar. Methods for Long-Term Electric Load Demand Forecasting; A Comprehensive Investigation. *IEEE International Conference on Industrial Technology*, 2008. ICIT 2008. 21-24 April 2008.

[2] K.A. Clents, A. J. Wood, Error Analysis of Econometric Load Forecasting Models. *IEEE Transactions on Power Apparatus and Systems*, Vol.PAS-98, No.2 March/April 1979.

[3] R.J. Wonnacott, T.H. Wonnacott, Econometrics, John Wiley and Sons, New York, 1975.

[4] T. Q. D. Khoa, L. M. Phuong, P.T.T.B inh, N. T. H. Lien. "Application of wavelet and neural network to long-term load forecasting," *International conference on power system technology* (POWERCON 2004), Singapore, 21-24 November 2004, pp. 840-844.

[5] Khaled M. EL_Naggar, and Khaled A. AL-Rumaih, "Electric load forecasting using genetic based algorithm, optimal filter estimator and least error square technique: Comparative study," *Transaction of engineering, computing and technology* , Vol.6, June 2005, ISSN 1305- 5313, pp. 138-142.

[6] M. A. Faraht, "Long-term industrial load forecasting and planning using neural networks technique and fuzzy interface method," *39th International universities power engineering conference*, UPEC 2004, vol. 1, pp. 368-372.

[7] J.H. Broehl. An end-use approach to demand forecasting. *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-100, No. 6 June 1981.

[8] M.Taradar Heque, and A.M.Kashtiban, "Application of neural networks in power systems; A review," *Transaction of engineering, computing and technology,* Vol.6, June 2005, ISSN 1305-5313, pp. 53-57.

[9] Amir F. Atiya, "Development of an intelligent long-term electric load forecasting system," *Proceedings of the International Conference*, ISAP apos, 1996, pp. 288-292.

[10] Bahman S. Kermanshahi, and Hiroshi Iwamiya, "Up to year 2020 load forecasting using neural nets," *Electric power system research (Elsevier)*, vol. 24, No.9, 2002 ,pp. 789-797.