Electrical Power and Energy Systems 64 (2015) 311–327

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes

A hybrid model based on data preprocessing for electrical power forecasting

Liye Xiao [a], Jianzhou Wang [b,*], Xuesong Yang [a], Liyang Xiao [c]

[a] School of Physical Electronics, University of Electronic and Technology of China, Chengdu, China
[b] School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China
[c] Complex Systems in Interaction, University of Technology of Compiègne, Compiègne, France

ARTICLE INFO

Article history:
Received 11 April 2014
Received in revised form 2 July 2014
Accepted 6 July 2014
Available online 6 August 2014

Keywords:
Electrical power forecasting
Hybrid model
Forecasting accuracy
Model selection

ABSTRACT

Electrical power forecasting plays a vital role in power system administration and planning. Inaccurate forecasting can lead to the waste of scarce energy resources, electricity shortages, and even power grid collapses. On the other hand, accurate electricity power forecasting can enable reliable guidance for the planning of power production and the operation of a power system, which is also important for the continued development of the electrical power industry. Although thousands of scientific papers address electricity power forecasting each year, only a small number are devoted to developing a general model for electricity power prediction that improves performance in different cases. This paper proposes a hybrid forecasting model for electrical power prediction that incorporates several artificial neural networks and model selection. To evaluate the forecasting performance of the proposed model, this paper uses half-hourly electrical power data of the State of Victoria and New South Wales of Australia as a case study. The experimental results clearly indicate that for this particular dataset, the forecasting performance of the proposed hybrid model is outstanding compared to that of the single forecasting model.

Introduction

Electrical power forecasting plays an important role in the secure and economic operation of power systems, which has significant economic consequences. For example, such forecasting is used in the development of preventive maintenance plans, which are implemented to safeguard generators, for the reliable evaluation of power systems, and for scheduling dispatch. In addition, many important decisions are made based on the forecasted electric load [1]. Providing high accuracy power forecasts allows for risks reduction, thereby improving the economic and social benefits of power grid management, which reduces generation costs, improves the security of the power system for power systems and helps administrators develop an optimal plan of action [2]. Moreover, accurate forecasting loads are crucial data for forecasting the electrical power prices in the operation of electrical power markets [3]. Therefore, developing electrical power forecasting techniques to achieve higher accuracy is highly desirable. However, load forecasting in a given region remains a difficult and challenging task because electric power is unavoidably affected by many factors of a stochastic nature, such as economic development, regional industrial production, holiday periods, weather conditions, social change, electricity price, and population [4]. To date, many studies have examined power prediction in terms of the improvement of accuracy using various common methods, such as fuzzy inference, artificial neural network, expert systems, auto regressive integrated moving average (ARIMA) model, regression models, and hybrid algorithms [5–7]. For example, Pai and Hong [8] forecasted electrical load in Taiwan by applying support vector machines using a simulated annealing algorithm (SVMSA), which was determined to be superior to the single general regression neural networks model (GRNN) and the autoregressive integrated moving average (ARIMA) model based on the experiential results. Wang et al. [9] proposed a model based on a trend with seasonal adjustments in combination with the ε-SVR for short-term forecasting. Hong [10] developed an electrical power forecasting model that combined the seasonal recurrent support vector regression model with a chaotic artificial bee colony algorithm (SR-SVR-CABC), which produced more accurate forecasting results than those afforded by the TF-e-SVR-SA and ARIMA models. A knowledge-based expert system (ES) was implemented by Kandil et al. [11] to support the selection of the most suitable model for load forecasting, and the availability of the selected method was demonstrated by a practical application. Wang et al. [12] used a non-linear fractal extrapolation algorithm and the largest Lyapunov exponent for short-term load forecasting,

* Corresponding author. Tel.: +86 15339864602; fax: +86 411 84710484.
E-mail address: wjzdufe@gmail.com (J. Wang).

http://dx.doi.org/10.1016/j.ijepes.2014.07.029
0142-0615/

which is highly useful for the technical optimization of aviation economics and in fields that are related to dynamic processes. In general, the load forecasting approach can be grouped into two varieties: traditional methods, characterized by the use of time series, and modern intelligent methods, characterized by the use of artificial neural networks. The traditional methods are mainly time series analysis methods based on mathematical statistics, including the Kalman filtering method [13], the regression Box–Jenkins' autoregressive integrated moving average (ARIMA) method [14], and the analysis method [15]. Papalexopoulos and Hesterberg [16] used a regression-based approach in short-term load forecasting. A smooth transition periodic autoregressive (STPAR) model was applied to short-term load forecasting by Amaral et al. [17]. To forecast short-term loads in the Iranian electricity market, a singular spectral analysis method was employed and improved upon, in the work of Afshar and Bigdeli [18]. Taylor [19] used triple seasonal methods for short-term load forecasting. Although the traditional methods have the advantages of being based on mature theory and technology and using simple algorithms, each of these methods is based on a linear analysis; as a result, these methods are unable to forecast nonlinear load series accurately.

Due to this limitation, in recent years, with the rapid development of intelligence techniques, many intelligent forecasting methods have been applied for electrical power forecasting. In 1990, Elman proposed a partial recurrent network model Elman neural network (ENN) [20], and Li et al. [21] developed a hybrid model involving ENN to forecast short-term loads. Hippert et al. [22] and Kandil et al. [23] present short-term load forecasting models by the application of artificial neural networks (ANN) to model load dynamics. Liu et al. [24] used a genetic-algorithm-optimized back propagation neural network (GA-BP) for industrial load forecasting. Hsu and Chen [25] established an artificial neural network (ANN) based on the collection of empirical data to predict the peak load of Taiwan's region. Bakirtzis et al. [26] proposed a neural network short term load forecasting model for the Greek power system. Specht [27] developed the generalized regression neural network (GRNN), which is a type of strong regression tool and a probability neural network with a trends network structure. Due to its powerful non-linear mapping capability, robustness, and high fault tolerance and the simplicity of its network structure, the GRNN was effectively used for short-term load forecasting [28].

In general, it is rare that a single forecasting model is optimal in all cases. Each model has its own particular strengths and weaknesses [29]. When multiple forecasting models are available, as inspired by the excellent performances of the various hybrid models outlined above, instead of selecting the "best" model, consideration is given to the development of a hybrid forecasting method, which is regarded as an outstanding approach for taking advantage of the strengths of each model. This method follows a regression framework in which the actual values and individual forecasts are treated as the responses and predictors, respectively [30,31]. The forecasts of the hybrid models often have smaller out-of sample errors than those of each of the component models [32]. However, the properties of the individual forecast error may vary over time. For example, Amjady [33] built a new hybrid model for short-term load forecasting of power systems because a single model may not be sufficient to confirm all of the characteristics of a time-series. Khashei and Bijari [34] believed that the combination of such models stems from the hypothesis that neither one can ensure the true course of data generation. Zhang [35] suggested a hybrid method involving both artificial neural network (ANN) models and the auto-regressive integrated moving average (ARIMA) model. To improve the forecasting performance, Niu et al. [36] proposed a hybrid ANN model by combining some statistical methods. Fan and Chen [37] presented an adaptive hybrid model that based on an adaptive two-stage hybrid network with self-organized map (SOM) and support vector machine (SVM). Valenzuela et al. [38] presented a hybrid intelligent method involving fuzzy systems, evolutionary algorithms, and ANNs, and the results indicated that the developed hybrid ARIMA–ANN model was superior to the individual models when they were separately applied. Zhang [39] proposed a hybrid model that combines both ARIMA and ANN models to take advantage of the unique strength of the two models in linear and nonlinear modeling. The experimental results with real data sets indicated that the combined model can be an effective way to improve forecasting accuracy achieved by either of the models used separately. Che and Wang [40] developed a hybrid model known as SVRARIMA that integrates both the support vector regression (SVR) and auto-regressive integrated moving average (ARIMA) models; this hybrid model takes advantage of the SVR and ARIMA models' individual strengths in the field of nonlinear and linear modeling, respectively. A hybrid model for short-term load forecasting that integrates artificial neural networks and fuzzy expert systems was presented by Kim et al. [41] and the proposed hybrid model provided good forecasting accuracy of the mean absolute percentage errors below 1.3%. Huang et al. [42] proposed a new particle swarm optimization (PSO) approach to identify the autoregressive moving average with exogenous variable (ARMAX) model for one-day to one-week ahead hourly load forecasts. In a recent paper, Zhang et al. [43] predicted New South Wales's electricity prices from the national electricity market in Australia using a hybrid method that combined the auto-regressive integrated moving average (ARIMA) model, least squares support vector machines (LSSVM) and the wavelet transform. Moghram and Rahman [44] reviewed five forecasting methods, i.e., the multiple linear regression method, the time series method, the general exponential smoothing method, the state space and Kalman filter method, and the knowledge-based approach, for short-term load forecasting. The authors used each of the models to predict summer and winter loads separately, but none of the methods was considered to be distinctive. The transfer function model provided the best results over the summer months, whereas it resulted in the second worst accuracy over the winter months [45]. As a result, because a single forecasting model cannot always be optimal in any case, in this paper, we developed a hybrid model that combines BPNN, ENN, GRNN, and GA-BPNN. Through experiments and the Diebold–Mariano predictive accuracy test (DM test), the proposed hybrid model was demonstrated to be highly effective for load forecasting by providing high accuracy and stability all of the time, thereby overcoming the problems of a single model exhibiting an excessive calculation time and instability and yielding low forecasting accuracy in certain situations.

Electrical power forecasting model

There are many neural network models that are used for electrical load forecasting, including several classical prediction models and several prediction models that are based on the development and optimization of some classic models; herein, we choose three classical prediction models: the BP neural network, GRNN, the Elman neural network and an optimized model GABP neural network.

BP neural network

The BP neural network (BPNN) is a type of artificial neural network (ANN) that is trained using the BP algorithm. The BPNN is based on a mature theory and has been applied widely. This particular network has three layers: an input layer, a hidden layer and an output layer. Fig. 1 Part a shows the structure of the BP neural