Electrical Power and Energy Systems 62 (2014) 816–824

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

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

Techniques of applying wavelet de-noising into a combined model for short-term load forecasting

Jujie Wang [a], Jianzhou Wang [b,*], Yaning Li [a], Suling Zhu [c], Jing Zhao [c]

[a] School of Science, Nanjing University of Posts and Telecommunications, Nanjing 210046, China
[b] School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China
[c] School of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, China

ARTICLE INFO

Article history:
Received 8 June 2013
Received in revised form 30 April 2014
Accepted 13 May 2014
Available online 18 June 2014

Keywords:
STLF
Wavelet de-noising
SARIMA
BPNN
Combined model
WDCM

ABSTRACT

Electrical load time series are non-stationary and highly noisy because a variety of factors affect electrical markets. The direct forecasting of electrical load with noisy data is usually subject to large errors. This paper proposes a novel approach for short-term load forecasting (STLF) by applying wavelet de-noising in a combined model that is a hybrid of the seasonal autoregressive integrated moving average model (SARIMA) and neural networks. The process of the proposed approach first decomposes the historical data into an approximate part associated with low frequency and a detailed part associated with high frequencies via a wavelet transform. A SARIMA and a back propagation neural network (BPNN) are then established by the low-frequency signal to forecast the future value. Finally, the short-term load is forecasted by combining the prediction values of SARIMA and BPNN, and the weights of the combination are determined using a variance–covariance approach. To evaluate the performance of the proposed approach, the electricity load data in New South Wales, Australia, are used as an illustrative example. A comparison of the results with other models shows that the proposed model can effectively improve the forecasting accuracy.

Introduction

Electric load forecasting is an essential subject in electric power system operations and planning. Forecasts are needed for a variety of utility activities, such as generation scheduling, the scheduling of fuel purchases, maintenance scheduling and security analysis. Unfortunately, the electric load is essentially dynamic, nonparametric, and chaotic in nature. In fact, many factors affect the electric load in a region, such as the population, economic development, social change, weather conditions and industrial productions in the region, electricity price, and holiday periods. This influence implies that accurate load forecasting is not only of great interest but also extremely challenging to power system administrators.

Various load-forecasting techniques have been proposed and successfully applied to predict the different classes of power system load requirements in the past two decades. Generally, these methods used in the literature can be divided into two categories: statistical models and artificial intelligence (AI) models. Statistical models are identical to the direct random time-series model, including linear regression, the exponential smoothing method, and autoregressive integrated moving average (ARIMA) models [1–3]. They do well in short-term forecasts, and have been tested in short-term forecasting [4,5]. However, they are not perfect in forecasting. First, most of statistical models assume that the electric loads data is normally distributed, but it is well known that electric loads series is not a normally distributed. Second, the intermittent and stochastic characteristic of electric loads series need more complex functions for capturing the nonlinear relations, but most of these models are based on the assumption that a linear correlation structure exists among time series values. To overcome these limitations, many AI approaches have been proposed to address these problems. These AI approaches, which primarily include neural networks [6–19], expert system-based methods [20–22], fuzzy logic-based approaches [23,24], and genetic algorithms [25], have yielded impressive results in dealing with electric load prediction. However, electric load series often contain both linear and nonlinear patterns. Many research efforts have indicated that prediction methods depend on the data patterns and there is no single best prediction method that can be applied to any data patterns [26,27]. Therefore, combining different models can increase the chance of capturing different patterns in the data and improve the forecasting performance. This approach has led to the rapid development of hybrid models based on popular

* Corresponding author. Tel.: +86 931 8914050; fax: +86 931 8912481.
E-mail address: wjz@lzu.edu.cn (J. Wang).

http://dx.doi.org/10.1016/j.ijepes.2014.05.038
0142-0615/

methods. For instance, Sharaf and Tjing [28] proposed a novel neuro-fuzzy short-term load forecasting model based on neural networks and fuzzy logic. Huang et al. [29] proposed a Grey-Markov forecasting model to forecast the electric power demand in China. Pao [30] applied hybrid non-linear models for energy consumption forecasting in Taiwan. Zhang and Dong [31] proposed an adaptive neural-wavelet model for short-term load forecasting in the competitive electricity market, and El-Keib et al. [32] and Maia and Goncalves [33] applied a hybrid model for load forecasting.

The wavelet transform can effectively remove the useless information in a time series. The seasonal autoregressive integrated moving average model (SARIMA) mainly addresses linear relationships, while the back propagation neural network model (BPNN) can handle nonlinear patterns. Considering the actual features of power load, a wavelet de-noising-based combined model (WDCM) for short-term electrical load forecasting is proposed in this paper by applying a wavelet transform to a hybrid model. This hybrid model consists of SARIMA and BPNN. In this model, the original data are first decomposed into an approximate part associated with a low frequency and a detailed part associated with high frequencies using the wavelet transform. SARIMA and BPNN are then established using the low-frequency signal to forecast the future value. Finally, the short-term load is forecast by combining the predicted values of SARIMA and BPNN, and the weights of the combination are determined using the "variance–covariance" approach. To evaluate the performance of the proposed approach, the electricity load data in New South Wales, Australia, were used as an illustrative example. A comparison of the results with the individual models and the basic combined model shows that the proposed model can effectively improve the forecasting accuracy.

The remaining sections of this paper are organized as follows. 'Individual forecasting models used in the combined model' introduces the essential methods of wavelet de-noising, SARIMA, and BPNN. The hybrid methodology is introduced in 'The wavelet de-noising-based combined model (WDCM)'. 'Experimentation design and results' presents the experimentation design and results. Finally, concluding remarks and future work are given in 'Conclusions and future work'.

Individual forecasting models used in the combined model

Wavelet transform

A wavelet transform is used to analyze the non-stationary time series in order to generate information on both the time and frequency domains. This transform may be regarded as a special type of Fourier transform at multiple scales that decomposes a signal into shifted and scaled versions of a "mother" wavelet. The continuous wavelet transform, denoted by CWT, is defined as the convolution of a time series $x(t)$ with a wavelet function $w(t)$ [34]:

$$CWT_x^{\psi}(b,a) = \varphi_x^{\psi}(b,a) = \frac{1}{\sqrt{|a|}} \int x(t) \cdot \psi^*\left(\frac{t-b}{a}\right) dt \quad (1)$$

where a is a scale parameter, b is the translational parameter and $*$ is the complex conjugate of $\psi(t)$. Let $a = 1/2^s$ and $b = k/2^s$, where s and k belong to the integer set Z. The CWT of $x(t)$ is a number at $(k/2^s, 1/2^s)$ on the time-scale plane. It represents the correlation between $x(t)$ and $\psi^*(t)$ at that time-scale point. A discrete version of Eq. (1) is thus obtained as

$$DWT_x^{\psi}(k,s) = \phi_x^{\psi}\left(\frac{k}{2^s}, \frac{1}{2^s}\right) = \int_{\infty}^{\infty} x(t) \cdot \psi \cdot \left(\frac{t-k/2^s}{1/2^s}\right) dt \quad (2)$$

which separates the signal into components at various scales corresponding to successive frequencies.Note that the DWT corresponds to the multi-resolution approximation expressions for the analysis of a signal in many frequency bands (or at many scales). In practice, multi-resolution analysis is carried out by starting with two channel filter banks composed of a low-pass and a high-pass filter, and each filter bank is then sampled at a half rate of the previous frequency. The number of steps of this de-composition procedure will depend on the length of the data. The down sampling procedure maintains the scaling parameter constant $(1/2)$ throughout successive wavelet transforms [35].

Over the past decade, DWT has been well developed and applied to analyze the signals in various fields [36]. In this study, DWT is utilized to remove noise from the electric load data for prediction purposes.

Seasonal autoregressive integrated moving average model (SARIMA)

Introduced by Box and Jenkins [37,38], the SARIMA model, which originates from the autoregressive model (AR), the moving average model (MA) and the autoregressive and moving average model (ARMA) models [39,40], is a classical statistical forecasting tool and has become one of the most popular models for time series forecasting analysis. Because the application of this model is very common, it is only briefly described here.The mean value of the time series $\{x_t | t = 1,2,\ldots,k\}$ is assumed to be zero. A non-seasonal ARIMA model of order (p,d,q) (denoted by ARIMA (p,d,q)) representing the time series can be expressed as

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \quad (3)$$

or

$$\phi(B)\nabla^d x_t = \theta(B) e_t \quad (4)$$

where x_t and e_t are the actual value and random error at time period t, respectively; $\phi_i (i = 1,2,\ldots,p)$ is a finite set of parameters and is determined via a linear regression; $\theta_j (j = 1,2,\ldots,q)$ is a finite set of weight parameters; p is an integer and often referred to as orders of the autoregressive; q is an integer and often referred to as orders of the moving average; B denotes the backward shift operator, $\nabla^d = (1-B)^d$; d is the order of regular differences; $\phi(B)$ and $\theta(B)$ are defined as $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ and $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$, respectively.

Particularly, e_t is assumed to be an independent and identically distributed normal random variable with mean zero and variance σ^2, and the roots of $\phi(x) = 0$ and $\theta(x)$ all lie outside the unit circle [41]. Similarly, a seasonal ARIMA (SARIMA) model can be written as follows (using the second expression):

$$\phi(B)\psi(B^s)\nabla^d(1-B^s)^D x_t = \theta(B)\Theta(B^s) e_t \quad (5)$$

where $\psi(B^s) = 1 - \psi_1 B^s - \psi_2 B^{2s} - \cdots - \psi_p B^{ps}, \theta(B^s) = 1 - \theta_1 B^s - \theta_2 B^{2s} - \cdots - \theta_q B^{qs}$; D is the number of seasonal differences, and s is the period. If the time series mean is $\mu \neq 0$, we replace x_t with $x_t - \mu$.

The choice of p, d, q, and s is very important in the SARIMA model building process, which is typically repeated several times until a satisfactory model is finally selected. The final selected model can then be used for prediction purposes. In this paper, we selected $p = 2$, $d = 0$, $q = 1$, and $s = 24$.

Back propagation neural network (BPNN)

When the linear restriction of the model form is relaxed, the possible number of nonlinear structures that can be used to describe and forecast a time series is enormous. Artificial neural networks are one of these models that can approximate various nonlinearities in the data.