Computers & Industrial Engineering 63 (2012) 695–701

Contents lists available at SciVerse ScienceDirect

Computers & Industrial Engineering

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

# Day-ahead electricity price forecasting by a new hybrid method

Jinliang Zhang *, Zhongfu Tan, Shuxia Yang

*Institute of Electric Power Economics, North China Electric Power University, Beijing 102206, PR, China*

ARTICLE INFO

*Article history:*
Available online 7 April 2012

*Keywords:*
Electricity price forecasting
WT
ARIMA
LSSVM
PSO

ABSTRACT

Electricity price forecasting has become necessary for power producers and consumers in the current deregulated electricity markets. Seeking for more accurate price forecasting techniques, this paper proposes a new hybrid method based on wavelet transform (WT), autoregressive integrated moving average (ARIMA) and least squares support vector machine (LSSVM) optimized by particle swarm optimization (PSO) to predict electricity prices. The proposed method is examined by using the data from New South Wales (NSW) of Australian national electricity market. Empirical testing indicates that the proposed method can provide more accurate and effective results than the other price forecasting methods.



## 1. Introduction

The deregulation of electric power industry has become a crucial issue around the world. The main objective of deregulation is to increase efficiency through competition. In the new environment, there are two major ways for electricity trading: bilateral contracts and the pool market. This paper considers itself in a pool-based electric energy market, because it is the most common arrangement in practice. In such a market, the producers and consumers submit their own bids that consist of a set of quantities with their prices. Then the market operator uses a market clearing algorithm to determine the prices. In this way, the prediction of electricity price is important to producers to maximize their profits and to consumers to maximize their utilities, respectively. Therefore, this paper is focused on the day-ahead market clearing price (MCP) forecasting. However, due to the complicated factors affecting electricity prices, the price series presents a complex behavior, which makes forecasting very challenging. Thus, a good price prediction method should be able to capture the complex behavior associated with price series.

In recent years, many methods have been proposed for short-term electricity price forecasting. Among these methods, two widely used approaches are time series models and artificial neural network (ANN). Time series models such as dynamic regression (DR) and transfer function (TF) (Nogales, Contreras, Conejo, & Espinola, 2002), ARIMA (Contreras, Espinola, Nogales, & Conejo, 2003), generalized auto-regressive conditional heteroskedastic (GARCH) (Garcia, Contreras, Akkeren, & Garcia, 2005), ARIMA-EGARCH (Bowden & Payne, 2008), GIGARCH (Diongue, Guegan, & Vignal, 2009) have been proposed for this purpose. However, most of the time series models are linear predictors, which have difficulties in forecasting the hard nonlinear behavior of electricity price series (Amjady & Hemmati, 2006).

To solve this problem, ANN has been proposed for price forecasting, which is an effective way to deal with the complex nonlinear problem. Zhang, Luh, and Kasiviswanathan (2003) presented a cascaded neural-network (NN) structure for MCP prediction. Guo and Luh (2004) proposed the cascaded architecture of multiple ANN to forecast MCP. Zhang and Luh (2005) used a kind of extended Kalman filter combined with ANN to predict MCP. Very recent related papers have been considered by Vahidinasab, Jadid, and Kazemi (2008) and Areekul, Senjyu, Toyama, and Yona (2010), and so on. Although the main advantage of ANN is its nonlinear modeling capability, it has the weakness of locally optimal solution. To overcome this shortcoming, LSSVM is applied for electricity price forecasting. The reason behind the choice of LSSVM is its high accuracy and global solution. However, directly applying LSSVM model does not produce a better result. The reason is that the selection of the parameters in a LSSVM model has a heavy impact on the forecasting accuracy. Instead of using genetic algorithm (GA), PSO is used to optimize the LSSVM parameters, namely PLSSVM. Compared to PSO, GA is lack of knowledge memory, which will lead to time consuming and inefficiency in the searching suitable parameters of a LSSVM model.

Since it is difficult to completely know the features of electricity prices, hybrid method is proposed for price forecasting. The basic idea of different models combination in forecasting is to use each model's unique feature to capture different patterns in the data, which is an effective and efficient way to improve forecasts (Zhang, 2003). Combining Probability Neural Network (PNN) and Orthogonal Experimental Design (OED), an Enhanced PNN was proposed by Lin, Gow, and Tsai (2010). Tan, Zhang, Wang, and Xu (2010) proposed a novel hybrid method based on wavelet transform

* Corresponding author. Tel.: +86 010 51963749; fax: +86 010 80796904.
*E-mail address:* zhangjinliang1213@163.com (J. Zhang).

0360-8352/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cie.2012.03.016

combined with ARIMA and GARCH models. Unsihuay-Vila, Zambroni de Souza, Marangon-Lima, and Balestrassi (2010) presented a new hybrid approach based on nonlinear chaotic dynamics and evolutionary strategy to forecast electricity prices. Catalão, Pousinho, and Mendes (2011) proposed a hybrid approach, which was based on wavelet transform, neural networks and fuzzy logic. A new prediction strategy composed of probabilistic neural network and hybrid neuro-evolutionary system was presented by Amjady and Keynia (2011). It is reasonable to consider the electricity price series to be composed of a linear autocorrelation structure and a nonlinear component. Thus, a hybrid method that has both linear and nonlinear modeling capabilities can be a good strategy for price forecasting. Electricity price forecasting is difficult because unlike load, electricity price series present such features as non-constant mean and variance, high frequency. Thus, wavelet transform is used to convert the original price series into a set of constitutive series, which present a better behavior than the original price series. Therefore, they can be predicted more accurately. Hence, a hybrid method using WT, ARIMA and PLSSVM is proposed, where the ARIMA model captures the linear component and the PLSSVM model captures the nonlinear component. The motivation to adopt such a hybrid method is to use other method's unique feature to capture different patterns in the electricity price series. Both theoretical and empirical findings suggest that combining different methods can be an effective and efficient way to improve forecasts. The main contribution of this paper can be summarized as follows:

(1) Wavelet transform is used to convert the ill-behaved price series into a set of constitutive series. Then each subseries is separately predicted by a new hybrid model based on ARIMA and PLSSVM.
(2) A new hybrid prediction method combined with WT, ARIMA and PLSSVM is proposed in this paper for day-ahead electricity price forecasting. To the best of our knowledge, this method has never been presented in the literature.
(3) The proposed approach is compared with some other approaches to demonstrate its effectiveness regarding forecasting accuracy.

The rest of this paper is organized as follows. Section 2 provides a description of WT, ARIMA and PLSSVM models. The hybrid methodology is given in Section 3. Numerical results are presented in Section 4. Section 5 gives the conclusions of this paper.

## 2. Theoretical background of WT, ARIMA and PLSSVM models

### 2.1. Wavelet transform

A brief description of wavelet transform (DWT) is given in the following. The discrete wavelet transform is defined as:

$$W_{(m,k)} = 2^{-(m/2)}\sum_{t=0}^{T-1} f_{(t)}\phi_{(t-n\cdot 2^m/2^m)} \tag{1}$$

where $T$ is the length of the signal $f_{(t)}$. The scaling and translation parameters are functions of the integer variables $m$ and $k$ ($a = 2^m$, and $b = k \cdot 2^m$). $t$ is the discrete time index.

A fast algorithm to implement DWT has been used. This algorithm has two stages: decomposition and reconstruction. In the first stage, the original signal is decomposed into one approximation series and some detail series. The basic concept in this stage begins with the selection of mother wavelet and the number of decomposition levels. In this paper, a wavelet function of Daubechies wavelet of order 4 (DB4) is used. The reason is that this wavelet offers an appropriate trade-off between wave-length and smoothness, resulting in an appropriate behavior for price forecasts (Amjady & Keynia, 2008). Also, three decomposition levels are considered, since it describes the price series in a more thorough and meaningful way than the others (Amjady & Keynia, 2008). In the second stage, DWT components can be assembled back into the original signal. More details of this model can be found in Appendix A.

### 2.2. ARIMA model

The ARIMA model is widely used in the areas of non-stationary time series forecasting, which can be written as:

$$\varphi(B)(1-B)^d X_t = \theta(B)\varepsilon_t \tag{2}$$

where $X_t$ represents a non-stationary time series at time $t$, $\varepsilon_t$ is a white noise, $d$ is the order of differencing, $B$ is a backward shift operator defined by $BX_t = X_{t-1}$, $\varphi(B)$ is the autoregressive operator defined as: $\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \ldots - \varphi_p B^p$, and $\theta(B)$ is the moving average operator defined as: $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q$.

Generally, the standard statistical methodology to construct an ARIMA model includes four phases: data preparation, model identification, parameter estimation and diagnostic checking. First, the autocorrelation function (ACF) and partial ACF (PACF) are considered to identify the ARIMA ($p$, $d$, $q$) model. Then, estimation of model coefficients is achieved by means of the maximum likelihood method. Furthermore, the hypothesis of the model is validated by Akaike Information Criterion (AIC). More details of this model can be found in Appendix B.

### 2.3. PLSSVM model

#### 2.3.1. LSSVM model

LSSVM is proposed by Suykens and Vandewalle (1999). The reason behind the choice of LSSVM is that the LSSVM regression algorithm achieves the global solution by solving a set of linear equations, which allows LSSVM to be faster than SVM (Iplikci, 2006). Given a training data set of $N$ points $\{x_i, y_i\}_{i=1}^N$ with input data $x_i \in R^n$ and output data $y_i \in R$, the decision function can be defined as:

$$y(x) = w^T\varphi(x) + b \tag{3}$$

In Eq. (3), $\varphi(x)$ denotes the nonlinear function that maps the input space to a high dimension feature space; $w$ denotes the weight vector; $b$ is the bias term.

For the function estimation problem, the structural risk minimization is used to formulate the following optimization problem:

$$\text{Minimize}: \quad \frac{1}{2}\|w\|^2 + \frac{1}{2}c\sum_{i=1}^n \varepsilon_i^2 \tag{4}$$

$$\text{Subject to}: y_i = w^T\varphi(x_i) + b + \varepsilon_i, \ i = 1, \ldots, N$$

In Eq. (4), $c$ represents the regularization constant, and $\varepsilon_i$ represents the training error.

To derive the solutions $w$ and $\varepsilon$, the Lagrange multipliers are introduced as following:

$$L(w, b, \varepsilon, a) = \frac{1}{2}\|w\|^2 + \frac{1}{2}c\sum_{i=1}^n \varepsilon_i^2 - \sum_{i=1}^n a_i[w^T\varphi(x_i) + b + \varepsilon_i - y_i] \tag{5}$$

In Eq. (5), $a_i$ is the introduced Lagrange multiplier.

According to the Karush–Khun–Tucker conditions, the finally result into the LSSVM model for function estimation can be described as:

$$f(x) = \sum_{i=1}^n a_i K(x, x_i) + b \tag{6}$$

In Eq. (6), the dot product $K(x, x_i)$ is known as the kernel function. This paper applied the radial basis function (RBF), because it is a