Energy Conversion and Management 51 (2010) 1911–1917

Contents lists available at ScienceDirect

Energy Conversion and Management

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

Short-term electricity prices forecasting based on support vector regression and Auto-regressive integrated moving average modeling

Jinxing Che, Jianzhou Wang *

School of Mathematics and Statistics, Lan-zhou University, Lan-zhou 730000, China

ARTICLE INFO

Article history:
Received 6 November 2008
Received in revised form 9 July 2009
Accepted 18 February 2010
Available online 29 March 2010

Keywords:
Support vector regression
ARIMA
Artificial neural networks
Price forecasting
Competitive market

ABSTRACT

In this paper, we present the use of different mathematical models to forecast electricity price under deregulated power. A successful prediction tool of electricity price can help both power producers and consumers plan their bidding strategies. Inspired by that the support vector regression (SVR) model, with the ε-insensitive loss function, admits of the residual within the boundary values of ε-tube, we propose a hybrid model that combines both SVR and Auto-regressive integrated moving average (ARIMA) models to take advantage of the unique strength of SVR and ARIMA models in nonlinear and linear modeling, which is called SVRARIMA. A nonlinear analysis of the time-series indicates the convenience of nonlinear modeling, the SVR is applied to capture the nonlinear patterns. ARIMA models have been successfully applied in solving the residuals regression estimation problems. The experimental results demonstrate that the model proposed outperforms the existing neural-network approaches, the traditional ARIMA models and other hybrid models based on the root mean square error and mean absolute percentage error.

1. Introduction

Accurate and efficient electricity price forecasting becomes more and more important for electricity markets. Electricity prices predictions are used for various purposes, such as speculation, derivative pricing, risk management and real option valuation. As accurate modeling of spot prices is the cornerstone of the optimal scheduling of physical assets and valuing of real options [1,2], intensive studies have been carried out to improve the precision.

Usually statistical based modeling techniques are used for electricity prices forecasting [3], which has advantages of simple feature and strong expansion ability, but fails to the forecasting of tremendous system's evolution serial, especially when electricity prices system changes a lot, and classical chaos appear in the system. The traditional forecasting method too simple to simulate the complex and fast change of the electricity prices system. Chaos theory, an important discovery in nonlinear dynamic system, came into being from 1960s, until 1980s, it developed to be a new study with special concept system and method frame [4]. Artificial neural network (ANN) is an effective way to solve the complex nonlinear mapping problem, which possesses excellent robustness and error-tolerance. Among the many existing tools, the ANN has received much attention because of its clear model, easy implementation and good performance in solving nonlinear problems, and this makes it suitable for modeling and forecasting of changing complex electricity system and electricity serial. In order to increase the forecasting accuracy, it has been performed using supervised neural learning techniques [5,6]; while some have used an ART-type neural network [7], fuzzy clustering method and ANN [8], marquardt algorithm and feedforward networks [9], fuzzy neural network [10], wavelet transform based approach and ANN [11] and an evolutionary algorithms coupled with ANN [12].

These studies usually use BP neural network model – error reverse transform neural network, which contains a great many parameters. These parameters are always judged by experience, so the model is hard to be established [13]. Also, it has been observed that while the neural network (NN) give small error for training patterns, the error for testing patterns is usually of a larger order [14]. However support vector regression, with the ε-insensitive loss function, can overcome this shortcoming and improve the generalization capability of network. Recently, a hybrid methodology has been proposed for stock price forecasting, which applied the ARIMA model in capturing the linear patterns and the residuals are modeled by the SVR [15]. Yet the linear and nonlinear patterns could interact, and it is difficult to decide which one is the dominant part; for this reason, modeling linear patterns using linear method will change the nonlinear patterns, and vice versa. Inspired by that SVR is the preferred model for nonlinear patterns and that, compared with NN method, it keeps the linear patterns undamaged, we propose a model that combines both SVR and ARIMA models to take advantage of the unique strength of SVR and ARIMA

* Corresponding author. Tel./fax: +86 931 3336112.
E-mail addresses: jinxingche@163.com (J. Che), chejx07@lzu.cn (J. Wang).

0196-8904/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.enconman.2010.02.023

models in nonlinear and linear modeling for short-term electricity prices forecasting, which is called SVRARIMA. The presented model is believed to greatly improve the prediction performance of the single ARIMA model and the single SVR model in forecasting electricity prices. The forecasting model adopts some tricks, such as following phase space reconstruction theory, utilizing G–P algorithm to calculate the saturation embedded dimension – the number of input level spot [16], applying ARIMA model to the residuals regression estimation problems. At last these tricks improve the generalization capability of SVR.

In this paper, the price forecasting of the California electricity market was used to examine the forecasting accuracy of the proposed hybrid model, the existing neural-network approaches, the traditional ARIMA models and other hybrid models. This study suggests that researchers and practitioners should carefully consider the nature and intended use of electricity prices data if choosing between neural networks, statistical methods and other hybrid models for electricity market management. The result of experiments proves that time serial forecasting and control system based on this method has the following advantages: (1) Better precision. (2) Smaller sampling variation effects. (3) Rather robust to parameter variation. (4) Faster response capability. (5) Compared with NN, it can keep the linear patterns undamaged relatively.

The remaining sections of this paper are organized as follows. In Section 2, the SVRARIMA model for forecasting is presented and the main steps of the method is given. Then, the possible reason behind the proposed technique is explained. In Section 3, the research design and the data description are outlined. Two performance measures are described. Obtained numerical results and comparisons are presented and discussed in Section 4. A brief review of this paper and the future research are in Section 5.

2. The proposed method

In this section, we describe the general theory of the proposed method. It can be found that SVR and ARIMA model will be effective in nonlinear and linear modeling separately. This paper investigates the effectiveness of a hybrid model for forecasting short-term electricity prices. Because of the ε-insensitive of SVR model and the minor linear patterns in electricity markets, this paper employs SVR to capture the nonlinear patterns firstly and ARIMA to regress the residuals.

2.1. Nonlinear dynamics

The goal of forecasting is to deduce objects' future according to their own development rules. According to Kolmogorov theorem, every time serial is a nonlinear input–output system whose states change with time in a determinstic way [17]. Accompanying with the development and maturity of neural network theory and phase space reconstruction technology, support vector regression is provided for nonlinear forecasting of electricity prices time serial.

Set disperse time serial $\{X(t)\}(t = 1, 2, \ldots, n)$, reconstruction phase space $\{Y(j)\}$, time-delays τ and proper embedded dimension $m(j = 1, 2, \ldots, N, N = n - (m - 1))$, and:

$$Y(j) = [X(j), X(j-\tau), X(j-2\tau), \ldots, X(j-(m-1)\tau)] \tag{2.1}$$

Judged by Takens embedding theorem, we have a mapping $\mathsf{F} : R^m \to R^m$ which makes $Y(j+\tau) = \mathsf{F}(Y(j))$ [18]. That is:

$$\begin{bmatrix} X(j-\tau) \\ X(j) \\ \vdots \\ X(j-(m-2)\tau) \end{bmatrix} = \mathsf{F}\left(\begin{bmatrix} X(j-\tau) \\ X(j) \\ \vdots \\ X(j-(m-2)\tau) \end{bmatrix}\right) \tag{2.2}$$

Then, in the reconstructed phase space there is a smooth map $\overline{\mathsf{F}} : R^m \to R$ such that:

$$X(j+\tau) = \overline{\mathsf{F}}(X(j), X(j-\tau), X(j-2\tau), \ldots, X(j-(m-1)\tau)) \tag{2.3}$$

Thus time series prediction can be carried out as long as a concrete function expression is found. But the function can be obtained by utilizing SVR's capability of approaching nonlinear reflection.

2.2. Support vector regression

Support Vector Machines (SVM) [19] are a set of classification and regression techniques, that are designed to systematically optimize its structure based on the input training data. This subsection briefly introduces support vector regression (SVR), which can be used for time-series forecasting; and for a more thorough coverage of SVM we refer the reader to the excellent surveys [20–22].

Given training data $(x_1, y_1), \ldots, (x_n, y_n)$, where x_i are the input vectors and y_i are the associated output values of x_i. In ε-SVR [23] the goal is to find a function $f(x)$ whose deviation from each target y_i is at most ε for all training data, and at the same time, is as "flat" as possible. Thus, the support vector regression is an optimization problem:

$$\min_{\omega, b, \xi, \xi^*} \quad \frac{1}{2}\omega^T\omega + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{2.4}$$

$$\text{subject to} \begin{cases} y_i - (\langle \omega, x_i \rangle + b) \leqslant \varepsilon + \xi_i \\ (\langle \omega, x_i \rangle + b) - y_i \leqslant \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geqslant 0 \end{cases} \tag{2.5}$$

where n denotes the number of samples, ξ_i represents the upper training error, and ξ_i^* is the lower training error subject to ε-insensitive tube $|y_i - (\langle \omega, x_i \rangle + b)| \leqslant \varepsilon$, whereas $C > 0$ is the regularized constant determining the trade-off between the empirical error and the regularization term. Notice that this tradeoff very robust to outliers, which makes SVR rather different from traditional error minimization problems. In other words, the SVR fits $f(x)$ to the data such that: (i) the training error is minimized by minimizing ξ_i and ξ_i^*, and (ii) $\frac{1}{2}\omega^T\omega$ is minimized to raise the flatness of $f(x)$, or to penalize excessively complex fitting functions. Moreover, we stress that the SVR problem can be seen as an *extension* of more traditional regression techniques: for instance, when $|\xi|^2$ is used as the ε-insensitive loss function, then we fall into the case of a minimum square error regression problem [24]. Finally, by introducing Lagrange multipliers, Kernel trick and exploiting the optimality constraints, the decision function has the following explicit form

$$f(x, \alpha_i, \alpha_i^*) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)K(x, x_i) + b \tag{2.6}$$

2.2.1. Lagrange multipliers

In Eq. (2.6), α_i and α_i^* are the so-called Lagrange multipliers. They are obtained by maximizing the dual function of Eq. (2.4), and the maximal dual function in Eq. (2.4) which has the following form:

$$\max_{\alpha_i, \alpha_i^*} \quad \sum_{i=1}^{n} y_i(\alpha_i - \alpha_i^*) - \varepsilon\sum_{i=1}^{n}(\alpha_i + \alpha_i^*) - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\alpha_i - \alpha_i^*)\left(\alpha_j - \alpha_j^*\right)\langle x_i, x_j \rangle \tag{2.7}$$

$$\text{subject to} \begin{cases} \sum_{i=1}^{n}(\alpha_i - \alpha_i^*) = 0 \\ 0 \leqslant \alpha_i \leqslant C \\ 0 \leqslant \alpha_i^* \leqslant C \\ i = 1, 2, \ldots, n \end{cases} \tag{2.8}$$