Omega 41 (2013) 517–524

Contents lists available at SciVerse ScienceDirect

Omega

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

# Carbon price forecasting with a novel hybrid ARIMA and least squares support vector machines methodology

Bangzhu Zhu [a,b,c,*], Yiming Wei [b,c]

[a] School of Economics and Management, Wuyi University, Jiangmen, Guangdong 529020, China
[b] School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China
[c] Center for Energy and Environmental Policy Research, Beijing Institute of Technology, Beijing 100081, China

ARTICLE INFO

*Article history:*
Received 8 September 2011
Accepted 22 June 2012
Available online 1 July 2012

*Keywords:*
Least squares support vector machine
ARIMA
Particle swarm optimization
Hybrid models
Time series forecasting
Carbon price

ABSTRACT

In general, due to inherently high complexity, carbon prices simultaneously contain linear and nonlinear patterns. Although the traditional autoregressive integrated moving average (ARIMA) model has been one of the most popular linear models in time series forecasting, the ARIMA model cannot capture nonlinear patterns. The least squares support vector machine (LSSVM), a novel neural network technique, has been successfully applied in solving nonlinear regression estimation problems. Therefore, we propose a novel hybrid methodology that exploits the unique strength of the ARIMA and LSSVM models in forecasting carbon prices. Additionally, particle swarm optimization (PSO) is used to find the optimal parameters of LSSVM in order to improve the prediction accuracy. For verification and testing, two main future carbon prices under the EU ETS were used to examine the forecasting ability of the proposed hybrid methodology. The empirical results obtained demonstrate the appeal of the proposed hybrid methodology for carbon price forecasting.



## 1. Introduction

Climate change has been a serious challenge in the last few decades. In order to reduce greenhouse gas emissions at the lowest overall cost, the European Union (EU) launched the EU Emissions Trading Scheme (EU ETS) covering around 12,000 installations in 25 countries and six major industrial sectors in 2005. As the first multinational atmospheric greenhouse gas cap-and-trade system, EU ETS is the largest carbon market in the world to date [1], and has proven to be not only an important tool for mankind to address climate change, but also a major choice for investors to decentralize their investment risks [2]. Therefore, the need for more accurate forecasts of carbon prices is driven by the desire to reduce risk and uncertainty.

Recently, although carbon price analysis has become one of the key issues concerned by many academic researchers and business practitioners [2–6], very little literature regarding carbon price forecasting can be found. In fact, carbon price changes over time, which means that it can be treated as a time series process. Therefore, carbon price forecasting is a kind of time series forecasting. During the past few decades, various approaches have been developed for time series forecasting, among which the autoregressive integrated moving average (ARIMA) model has been found to be one of the most popular models due to its statistical properties, as well as the well-known Box–Jenkins methodology in the modeling process. However, the ARIMA model is only a class of linear model and thus it can only capture linear patterns in a time series. Therefore, the ARIMA model cannot effectively capture nonlinear patterns hidden in a time series.

In order to overcome the limitations of the linear models and account for the nonlinear patterns existing in real cases, numerous nonlinear models have been proposed, among which the artificial neural network (ANN) has shown excellent nonlinear modeling capability. Although a large number of successful applications have shown that ANN has been successfully adopted in many forecasting fields [7–10], ANN suffers from some weaknesses, such as locally optimal solutions and over-fitting, which can make the forecasting precision unsatisfactory. In 1995, support vector machine (SVM), a novel ANN, was developed by Vapnik [11]. Established on the structural risk minimization (SRM) principle by minimizing an upper bound of the generalization error, SVM can result in resistance to the over-fitting problem [12]. However, SVM formulates the training process through quadratic programming, which can take much more time. In 1999, Suykens and his colleagues proposed a novel SVM known as least squares support vector machine (LSSVM) [13], which is able to solve linear problems quicker with a more straight forward approach. Until now, LSSVM has been successfully used in pattern recognition and nonlinear regression estimation problems. At the same time, to obtain the optimal LSSVM model, it

* Corresponding author at: School of Economics and Management, Wuyi University, Jiangmen, Guangdong 529020, China. Tel.: +86 750 3299331; fax: +86 750 3299335.
*E-mail address:* wpzbz@126.com (B. Zhu).

0305-0483/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.omega.2012.06.005

is important to choose a kernel function and determine the kernel parameters [14]. Therefore, we introduce particle swarm optimization (PSO) [15] to optimize the parameters of LSSVM in this study.

Although both LSSVM and ARIMA models have achieved success in their own linear or nonlinear domains, neither is suitable for all circumstances. The approximation of ARIMA models to complex, nonlinear problems, as well as LSSVM to model linear problems, may be inappropriate, let alone in problems that contain both linear and nonlinear patterns. Since it is difficult to completely know the characteristics of a real situation, a hybrid methodology that has both linear and nonlinear modeling capabilities can be a good candidate for practical use [16], which has been demonstrated by numerous studies [16–25]. However, existing hybrid methodologies often combine the traditional ANN [16–21] or SVM [22–25] and ARIMA models. Until now, no LSSVM and ARIMA hybrid model has been found for carbon price forecasting; therefore, this study aims to fill this gap.

The contributions made by this paper are two-fold. Firstly, we establish a novel LSSVM and ARIMA hybrid forecasting methodology to forecast carbon prices. In our proposed methodology, carbon prices are decomposed into two components: a linear component and a nonlinear component. An ARIMA model and a LSSVM model are used to capture the linear and nonlinear components of carbon prices respectively, and their forecasting values are integrated into the final forecasting results. Furthermore, PSO is used to find the optimal LSSVM parameter settings to forecast carbon prices in the future. Secondly, we evaluate the forecasting performance of the single ARIMA, ANN and LSSVM models, the hybrid ARIMA and ANN model, the hybrid ARIMA and SVM model, and the hybrid ARIMA and LSSVM model by forecasting two main future carbon prices under the EU ETS. The empirical results obtained demonstrate that the proposed hybrid model can outperform the single ARIMA, ANN, LSSVM models, the hybrid ARIMA and ANN model, and the hybrid ARIMA and SVM model.

The remainder of this study is organized as follows: Section 2 describes the ARIMA, LSSVM and hybrid models; Section 3 elaborates on the PSO–LSSVM model; Section 4 reports the experimental results and Section 5 provides the overall conclusions.

## 2. Methodology

### 2.1. ARIMA model

In the ARIMA model, carbon price is a linear function of past values and error terms. An ARIMA ($p$, $d$, $q$) model of degree of AR ($p$), difference ($d$) and MA ($q$) can be mathematically expressed as

$$x_t = u_t + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \cdots + \varphi_p x_{t-p} - \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

where $x_t$ is the carbon price obtained by differencing $d$ times, $\varepsilon_t$ (hypothetical white noise) is assumed to be independently and identically distributed with a mean of zero and a constant variance of $\sigma_\varepsilon^2$, $p$ and $q$ are the numbers of autoregressive and moving average terms in the ARIMA model, and $\varphi_i$ $(i=1,2,\ldots,p)$ and $\theta_i$ $(i=1,2,\ldots,q)$ are the model parameters to be estimated.

Fitting an ARIMA model to carbon price involves the following four-step iterative processes: determine the structure of ARIMA model, estimate the parameter values of the ARIMA model, perform ARIMA model tests on the residuals, and predict future carbon prices.

The major advantage of the ARIMA model is that it can capture the linear patterns of carbon prices well and is relatively easy to use. However, ARIMA on its own is not adequate for carbon price forecasting because real carbon prices are often nonlinear and irregular. Therefore, we introduce the LSSVM model to capture nonlinear patterns existing in the carbon price.

### 2.2. Least squares support vector machine for regression

In contrast to other forecasting approaches, SVM, firstly proposed by Vapnik [11] in 1995 and based on the SRM principle, has been successfully applied to classification and regression. However, SVM training is a time consuming process when analyzing huge data. LSSVM is a modification of the standard SVM developed by Suykens et al. [13] in 1999 to overcome these shortcomings, which results in a set of linear equations instead of a quadratic programming problem. Consider a given training set $x_i, y_i$, $i=1,2,\ldots,l$ with input data, $x_i$, and output data, $y_i$. LSSVM defines the regression function as

$$\min J(w,e) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{i=1}^{l} e_i^2 \tag{1}$$

subject to

$$y_i = w^T \phi(x_i) + b + e_i, \; i=1,2,\ldots,l \tag{2}$$

where $w$ is the weight vector, $\gamma$ is the penalty parameter, $e_i$ is the approximation error, $\phi(\cdot)$ is the nonlinear mapping function and $b$ is the bias term. The corresponding Lagrange function can be obtained by

$$L(w,e,\alpha,b) = J(w,e) - \sum_{i=1}^{l} \alpha_i w^T \phi(x_i) + b + e_i - y_i \tag{3}$$

where $\alpha_i$ is the Lagrange multiplier. Using the Karush–Kuhn–Tucker (KKT) conditions, the solutions can be obtained by partially differentiating with respect to $w$, $b$, $e_i$and $\alpha_i$:

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \to w = \sum_{i=1}^{l} \alpha_i \phi(x_i) \\ \frac{\partial L}{\partial b} = 0 \to \sum_{i=1}^{l} \alpha_i = 0 \\ \frac{\partial L}{\partial e_i} = 0 \to \alpha_i = \gamma e_i \\ \frac{\partial L}{\partial \alpha_i} = 0 \to w^T \phi(x_i) + b + e_i - y_i = 0 \end{cases} \tag{4}$$

By eliminating $w$ and $e_i$, the equations can be changed into

$$\begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 & I_v^T \\ I_v & \Omega + \gamma^{-1} I \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{5}$$

where $y=[y_1,y_2,\ldots,y_l]^T$, $I_v=[1,1,\ldots,1]^T$, $\alpha=[\alpha_1,\alpha_2,\ldots,\alpha_l]^T$, and the Mercer condition [11] has been applied to matrix $\Omega$ with $\Omega_{km} = \phi(x_k)^T \phi(x_m)$, $k,m=1,2,\ldots,l$. Therefore, the LSSVM for regression can be obtained from

$$y(x) = \sum_{i=1}^{l} \alpha_i K(x,x_i) + b \tag{6}$$

where $K(x,x_i)$ is the kernel function.

A major advantage of the LSSVM model is that it can capture the nonlinear patterns hidden in the carbon price. However, using the LSSVM model alone to model the carbon prices may produce mixed results [26,27]. Therefore, we can conclude that the relationship between the ARIMA and LSSVM models is complementary, and it is necessary to combine the two for effective carbon price forecasting.

### 2.3. The hybrid models

In reality, carbon prices are rarely purely linear or nonlinear, but often contain both linear and nonlinear patterns, due to their inherently high complexity. This makes carbon price forecasting