CRUDE OIL PRICE FORECASTING WITH AN IMPROVED MODEL BASED ON WAVELET TRANSFORM AND SUPPORT VECTOR MACHINES

Ruhaidah Samsudin¹, Ani Shabri²

¹Department of Software Engineering, Faculty of Computing, Universiti Teknologi Malaysia, 81310, Johor, Malaysia
ruhaidah@utm.my

²Mathematic Department, Faculty of Science, Universiti Teknologi Malaysia, 81310, Johor, Malaysia
ani@utm.my

Abstract

This paper presents a hybrid wavelet support vector machines (WSVM) model that combines both wavelet technique and the SVM model for crude oil price forecasting. Based on the purpose, the main time series was decomposed to some multi-frequently time series by wavelet theory and these time series were imposed as input data to the SVM for forecasting of crude oil price series. To assess the effectiveness of this model, daily crude oil market-West Texas Intermediate (WTI) has been used as the case study. Time series prediction capability performance of the WSVM model is compared with the single SVM model using various statistics measures. As seen in comparison, WSVM yielded more accurate than of any individual model and offered a practical solution to the problem in crude oil price forecasting.

Keywords: Wavelet, support vector machines, crude oil, forecasting

1. Introduction

Highly accurate and reliable crude oil forecasts is very important for component for business and governments around the world to make strategic for option pricing formulas, portfolio allocation and risk measurement. Over the past several decades, much effort has been devoted to the development and improvement of crude oil forecasting models.

In recent years, Computational Intelligence (CI) model is applied to solve problems for some physical processes with nonlinear relations. Most important models such as: Artificial neural network, fuzzy logic, genetic algorithm and a new one, called support vector machine (SVM) are used. SVM is one of the newest methods that have attracted many researchers in various scientific fields. SVMs developed by Vapnik (1998) have provided another novel approach to improve the generalisation property of neural networks. Originally, SVMs were developed for pattern recognition problems and regression estimation problems. Recently, SVM has been applied in various areas of time series forecasting and successful used for financial forecasting (Cao & Tay, 2001; Kim 2003; Perez-Cruz et al. 2003; Chen et al. 2010). In crude oil prices field, the SVM method has received very little attention and only a few applications of SVM to modeling of crude oil prices forecasting have been carried out (Khashman & Nwulu, 2011).

Wavelet transform has been studied for many years by mathematicians and has been receiving increasing attention in different areas of financial successfully (Yousefi et al. 2005). Recently, new hybrid models on wavelet transform process with the CI models have been improved for forecasting (Liu et al. 2007; Jammazi & Aloui 2012; Ortega 2012; Chao et al. 2012). They observed that use of wavelet techniques to pre-process time series data into decomposed wavelet coefficients of different decompositions produced significantly better results than original time series when used as inputs.
In this paper, a hybridization of wavelet and SVM model is proposed to forecast daily crude oil prices market-West Texas Intermediate (WTI). To verify the application of this model, the proposed model was compared with individual SVM time series model.

2 Support Vector Machines

The basic idea of SVM for function approximation is mapping the data \( x \) into a high-dimensional feature space by a nonlinear mapping and then performing a linear regression in the feature space. Consider a given training set of \( n \) data points \( \{x_i, y_i\}_{i=1}^n \) with input data \( x_i \in \mathbb{R}^p \), \( p \) is the total number of data patterns) and output \( y_i \in \mathbb{R} \). SVM approximate the function in the following form

\[
y(x) = w^T \varphi(x) + b
\]  

(1)

where \( \varphi(x) \) represents the high dimensional feature spaces, which is nonlinearly mapped from the input space \( x \). The coefficient are \( w \) and \( b \) are estimated by minimizing the regularized function

\[
R(C) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n L_{\varepsilon}(d_i, y_i)
\]  

(2)

\[
L_{\varepsilon}(d_i, y_i) = \begin{cases} 
|d_i - y_i| - \varepsilon, & |d_i - y_i| \geq \varepsilon \\
0, & \text{otherwise}
\end{cases}
\]  

(3)

By introducing Lagrange multipliers and exploiting the optimality constraints, the decision function given by Eq. (1) has the following explicit form:

\[
y(x) = \sum_{i=1}^n (a_i - a_i^\ast) K(x, x_i) + b
\]  

(4)

In Eq. (4), \( a_i \) and \( a_i^\ast \) are the so-called Lagrange multipliers. They satisfy the equalities \( a_i \ast a_i^\ast = 0 \), \( a_i \geq 0 \) and \( a_i^\ast \geq 0 \) where \( i = 1, 2, ..., n \). \( K(x, x_i) \) is defined as the kernel function. The value of the kernel is equal to the inner product of two vectors \( X_i \) and \( X_j \) in the feature space \( \varphi(x_i) \) and \( \varphi(x_j) \), that is, \( K(x_i, x_j) = \varphi(x_i) \ast \varphi(x_j) \). There are four basic kernel types currently in use with SVM; these are: the linear kernel, polynomial kernel, radial basis function (RBF) kernel, and the sigmoid kernel. In this work the RBF kernel is used and the equation for the RBF kernel is given

\[
K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \quad \gamma > 0
\]

Here \( \gamma \) is kernel parameter. The kernel parameter should be carefully chosen as it implicitly defines the structure of the high dimensional feature space \( \varphi(x) \) and thus controls the complexity of the final solution.
3 Wavelet Analysis

Nowadays wavelet-analysis is one of the most powerful tools in the study of time series. Wavelet transformations provide useful decomposition of original time series by capturing useful information on various decomposition levels. According to the Mallat’s theory, the original discrete time series \( x(t) \) can be decomposed into a series of linearity independent approximation and detail signals by using the inverse DWT. The inverse DWT is given by Mallat (1989)

\[
x(t) = T + \sum_{m=1}^{M} \sum_{n=0}^{2^{m-1}} W_{m,n} 2^{-m/2} \psi(2^{-m} t - n)
\]

or in a simple format as

\[
x(t) = A_M(t) + \sum_{m=1}^{M} D_m(t)
\]

which \( A_M(t) \) is called approximation sub-series or residual term at levels \( M \) and \( D_m(t) \) \((m = 1, 2, ..., M)\) are detail sub-series which can capture small features of interpretational value in the data.

4 An application

In this study, the West Texas Intermediate (WTI) crude oil price series was chosen as experimental sample. The daily data from January 1, 1986 to September 30, 2006, excluding public holidays, with a total of 5237 was employed as experimental data. For convenience of WLR modeling, the data from January 1, 1986 to December 31, 2000 is used for the training set (3800 observations), and the remainder is used as the testing set (1437 observations). Figure 1 shows the daily crude oil prices from January 1, 1986 to September 30. In practice, short-term forecasting results are more useful as they provide timely information for the correction of forecasting value. In this study, two performance criteria such as MSE and MAE were used to evaluate the accuracy of the models. These criteria are given below:

\[
\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{x}_i)^2 \quad \text{and} \quad \text{MAE} = \frac{1}{n} \sum_{j=1}^{n} |x_j - \hat{x}_j|
\]

where \( x_i \) is the actual value and \( \hat{x}_i \) is the forecasted value of period \( t \), and \( n \) is the number of total observation.

Fig. 1 The daily crude oil prices from January 1, 1986 to September 30, 2000
5.1 Fitting SVM to the data

For SVM learning, the RBF kernel was used. To make an efficient SVM model, three parameters $C$, $\varepsilon$ and $\gamma$ must be carefully predetermined. In this study, the ranges of $C$ was set to [1, 10] at increment of 1.0 and [0.1, 0.5] at increment of 0.1 for $\varepsilon$ and $\gamma$. The optimal values of $\{C,\varepsilon\}$ are selected using 10-fold cross-validation repeated ten times to increase the reliability of the results. For the final predictions used in the competition, we used all the available data to train the models with the parameters producing minimum validation errors. The parameters value that yields the minimum generalization error is, then, chosen.

5.2 Fitting Hybrid Wavelet LSSVM to the data

A hybrid wavelet-SVM (WSVM) model is obtained by combining two methods, DWT and SVM. Before SVM applications, the original time series data were decomposed into periodic components (DWs) by Mallat DWT algorithm (Mallat 1998). Three decomposition levels are employed in this study also same studies employed by Kisi (2010).

The effectiveness of wavelet components is determined using the determination coefficient ($R^2$) between the observed crude oil data and the wavelet coefficients of different decomposition levels (see Table 1). It is observed that 41.2% of the D1 component significant effect on crude oil data. The D2 and D3 show significantly higher $R^2$ compared to the D1 component. According to the $R^2$, the effective components D2 and D3 were selected as the dominant wavelet components and approximation (A3) component were added to each other to constitute the new series. For the WSVM model, the new series is used as inputs to the SVM model. Figure 2 shows the structure of the WSVM model. Figure 3 shows the original crude oil data time series and their Ds, that is the time series of 2-daily mode (D1), 4-daily mode (D2), 8-daily mode (D3), approximate mode (A3), and the combinations of effective details and approximation components mode (A2 + D2 + D3). Five different combinations of the new series input data is used for forecasting. The forecasting performances of the SVM and WSVM models is presented in Table 2 in terms of MSE and MAE.

![Input time series (e.g. previous time series)](image)

Decompose input using DWT

Output (e.g. current time series)

Add the effective Ds (details) and As (approximation) as input for SVM

SVM Model

Fig. 2 The structure of the WSVM model
Approximation (As) of L1 for the lags 1,2,3,4,5. Table 2 shows that WSVM model has a significant positive effect on crude oil forecast.

![Graphs of daily return, decomposed wavelet sub-series components (Ds) and Approximation (As) of crude oil prices WTI data.](image)

**Fig.3** Daily return, decomposed wavelet sub-series components (Ds) and Approximation (As) of crude oil prices WTI data

**Table 1** The determination coefficients ($R^2$) between each of sub-time series and original monthly Discrete Wavelets Components

<table>
<thead>
<tr>
<th>Discrete Wavelets</th>
<th>Coefficient of Determination ($R^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$D_{1,3}/Q_l$</td>
</tr>
<tr>
<td>D1</td>
<td>0.0040</td>
</tr>
<tr>
<td>D2</td>
<td>0.0154</td>
</tr>
<tr>
<td>D3</td>
<td>0.0730</td>
</tr>
<tr>
<td>A3</td>
<td>0.0139</td>
</tr>
</tbody>
</table>

**Table 2** Forecasting performance indices of SVM and WSVM

<table>
<thead>
<tr>
<th>Lags</th>
<th>Model</th>
<th>SVM</th>
<th>WSVM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MSE</td>
<td>MAE</td>
</tr>
<tr>
<td>I</td>
<td>M1</td>
<td>0.9397</td>
<td>0.7101</td>
</tr>
<tr>
<td>1.2</td>
<td>M2</td>
<td>0.9357</td>
<td>0.7140</td>
</tr>
<tr>
<td>1.2,3</td>
<td>M3</td>
<td><strong>0.9118</strong></td>
<td>0.7023</td>
</tr>
<tr>
<td>1.2,3,4</td>
<td>M4</td>
<td>0.9123</td>
<td><strong>0.6996</strong></td>
</tr>
<tr>
<td>1.2,3,4,5</td>
<td>M5</td>
<td>0.9257</td>
<td>0.7028</td>
</tr>
</tbody>
</table>

As seen from Table 2, SVM model with lags 3 (M3) has the smallest MSE (0.9118) and lags 4 (M4) has the smallest MAE (3.137). However, for the WSVM model, the best MSE (8.825) and MAE (1.952) were obtained for the lags 5 (M5). Table 2 shows that WSVM model has a significant positive effect on crude oil forecast.
Figure 4 shows the Box-plot for the SVM and WSVM models for forecasting period. It can be seen that the errors of WSVM model quite close to the zero. Overall, it can be concluded the WSVM model provided more accurate forecasting results than the other models for crude oil forecasting.

Figure 4 The Box-plot for the SVM and WSVM models for forecasting period

6 Conclusions
The new method based on the WSVM was developed by combining the discrete wavelet transforms (DWT) and support vector machines (SVM) model for forecasting crude oil prices. The daily crude oil prices time series data was decomposed at 3 decomposition levels (2–4–8 months). The sum of effective details and the approximation component were used as inputs to the SVM model. The performance of the proposed WSVM model was compared to regular SVM model. Comparison results indicated that the WSVM model was substantially more accurate than SVM model. The decomposed periodic components obtained from the DWT technique are found to be most effective in yielding accurate forecast when used as inputs in the SVM models. The accurate forecasting results indicate that WSVM model provides a superior alternative to SVM, and a potentially very useful new method for crude oil price forecasting.

Acknowledgment The authors thankfully acknowledged the financial support that afforded by MOHE and Universiti Teknologi Malaysia under GUP Grant (VOT 00K45).

References


