

## Research Article

# Crude Oil Price Forecasting Based on Hybridizing Wavelet Multiple Linear Regression Model, Particle Swarm Optimization Techniques, and Principal Component Analysis

Ani Shabri<sup>1</sup> and Ruhaidah Samsudin<sup>2</sup>

<sup>1</sup> Department of Science Mathematic, Faculty of Science, Universiti Teknologi Malaysia, 81310 Johor, Malaysia

<sup>2</sup> Department of Software Engineering, Faculty of Computing, Universiti Teknologi Malaysia (UTM), 81310 Johor, Malaysia

Correspondence should be addressed to Ani Shabri; [ani\\_sabri@hotmail.com](mailto:ani_sabri@hotmail.com)

Received 31 December 2013; Accepted 18 March 2014; Published 8 May 2014

Academic Editors: S. Balochian, V. Bhatnagar, and Y. Zhang

Copyright © 2014 A. Shabri and R. Samsudin. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Crude oil prices do play significant role in the global economy and are a key input into option pricing formulas, portfolio allocation, and risk measurement. In this paper, a hybrid model integrating wavelet and multiple linear regressions (MLR) is proposed for crude oil price forecasting. In this model, Mallat wavelet transform is first selected to decompose an original time series into several subseries with different scale. Then, the principal component analysis (PCA) is used in processing subseries data in MLR for crude oil price forecasting. The particle swarm optimization (PSO) is used to adopt the optimal parameters of the MLR model. 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 WMLR model is compared with the MLR, ARIMA, and GARCH models using various statistics measures. The experimental results show that the proposed model outperforms the individual models in forecasting of the crude oil prices series.

## 1. Introduction

Crude oil prices do play significant role in the global economy and constitute an important factor affecting government's plans and commercial sectors. Forecasting crude oil price is among the most important issues facing energy economists. Therefore, proactive knowledge of its future fluctuations can lead to better decisions in several managerial levels.

The literature dealing with forecasting crude oil is substantial. The application of the classical time series models such as autoregressive moving average (ARMA) (Yu et al. [1], Mohammadi and Su [2], and Ahmad [3]) and econometric model such as generalized autoregressive conditional heteroscedasticity (GARCH) type models (Agnolucci [4], Wei et al. [5], Liu and Wan [6]) for crude oil forecasting has received much attention in the last decade. But because the crude oil price has the volatility, nonlinearity, and irregularity, the classical and econometric model can lead to the decrease of the accuracy.

Due to the limitations of the classical and econometric models, soft-computing models, such as neural fuzzy (Ghafari and Zare [7]), artificial neural networks (Kaboudan [8], Mirmirani and Li [9], Shambora and Rossiter [10], and Yu et al. [11]), support vector machines (Xie et al. [12]), and genetic programming (GP), provide powerful solutions to nonlinear crude oil price prediction. Many experiments found that the soft-computing models often had some advantages over statistical-based models. However, these AI models also have their own shortcomings and disadvantages. For example, ANN often suffers from local minima and over-fitting, while other soft-computing models, such as SVM and GP, including ANN, are sensitive to parameter selection [1].

To remedy the above shortcomings, some hybrid methods have been used recently to predict crude oil price and obtain the best performances. In last year, wavelet transform has become a useful method for analyzing such as variations, periodicities, and trends in time series. The hybrid models with wavelet transform processes have been improved for

forecasting. For example wavelet-neural network (Jammazi and Aloui [13], Qunli et al. [14], and Yousefi et al. [15]), wavelet-least square support vector machines (LSVM) (Bao et al. [16]), and wavelet-fuzzy neural network (Liu et al. [17]) have been employed recently on some studies in crude oil forecasting. They observed that the wavelet transform fairly improves forecasting accuracy.

A major drawback of wavelet transform for direction prediction is that the input variables lie in a high-dimensional feature space depends on the number of sub-time series components. Because the number of sub-time series components for wavelet is inadvisable to be too many, in this study principal component analysis (PCA) is proposed to reduce the dimensions of sub-time series components.

The multiple linear regressions (MLR) model that is much easier to interpret is considered as an alternative to ANN model. In this paper, a hybrid wavelet multiple linear regression (WMLR) model integrating wavelet and MLR is proposed for short-term daily crude oil price forecasting. The study applies particle swarm optimization (PSO) to adopt the optimal parameters to construct the MLR model. For verification purpose, the West Texas Intermediate (WTI) crude oil spot price is used to test the effectiveness of the proposed WMLR ensemble learning methodology. Finally to evaluate the model ability, the proposed model was compared with individual ARIMA and GARCH models.

## 2. Methodology

**2.1. The ARIMA Model.** The most comprehensive of all popular and widely known statistical methods used for time series forecasting are Box-Jenkins models (Box and Jenkins [18]). It has achieved great success in both academic research and industrial applications during the last three decades. The general form of ARIMA models can be expressed as

$$y_t = \sum_{i=1}^p \phi_i y_{t-i} + \sum_{i=1}^q \theta_i e_{t-i} + e_t, \quad (1)$$

where  $p$  is the order of the autoregressive,  $q$  is the order of the moving average, and  $e_t$  is the random error. The Box-Jenkins methodology is basically divided into four steps: identification, estimation, diagnostic checking, and forecasting.

**2.2. The GARCH Model.** GARCH models have found extensive application in the literature and the most popular volatility model is GARCH (1, 1) model proposed by Bollerslev [19]. The standard GARCH (1, 1) can be described as follows:

$$r_t = \mu_t + \varepsilon_t = \mu_t + h_t^{1/2} \eta_t \quad \eta_t \sim N(0, 1), \quad (2)$$

$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1},$$

where  $\mu_t$  denote the conditional mean and  $h_t$  is the conditional variances and  $\eta_t$  is a standardized error and  $r_t = \ln(x_t/x_{t-1})$  is log return.

**2.3. Multiple Linear Regressions.** Multiple linear regressions (MLR) model is one of the modelling techniques to investigate the relationship between a dependent variable and several independent variables. Let the MLR have  $p$  independent variables with  $n$  observations. Thus the MLR can be written as

$$Y = w_0 + w_1 x_1 + w_2 x_2 + \dots + w_p x_p + \varepsilon_t, \quad (3)$$

where  $w$  are regression coefficients,  $Y$  is dependent variable,  $x_i$  are independent variables and  $\varepsilon_t$  is fitting errors. The method of least squares is generally used to estimate the coefficients model. In many applications, the results of a least squares fit are often unacceptable when the model is wrong or when the model is misspecified (Bozdogan and Howe [20]).

In this study, particle swarm optimization (PSO) method is presented to determine the optimal parameters of the MLR model. The PSO methods have proven to be very effective in solving a variety of difficult global optimization problems in forecasting (Chen and Kao [21] and Alwee et al. [22]), heat problem (Ma et al. [23] and Tyagi and Pandit [24]), and dynamic environments (Liu et al. [25]).

The classic solution of MLR model involves the minimization of the sum of the square errors between the model-predicted value and the corresponding data value:

$$\min f(w) = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2, \quad (4)$$

where  $n$  is the number of training data samples,  $Y_i$  is the actual value, and  $\hat{Y}_i$  is the forecasted value of train data. The same methodology was used to solve this problem using PSO algorithms. The solution with a smaller fitness  $f(w)$  of the training data set has a better chance of surviving in the successive generations.

**2.4. Particle Swarm Optimization.** Particle swarm optimization (PSO) is a population-based heuristic method inspired by the collective motion of biological organisms, such as bird flocking and fish schooling, to simulate the seeking behavior to a food source (Bratton and Kennedy [26]). The population of PSO is called a swarm and each individual in the population of PSO is called a particle. The PSO begins with a random population and searches for fitness optimum just like genetic algorithm (GA). To find the optimum solution, each particle adjusts the direction through the best experience which it has found ( $p_{\text{best}}$ ) and the best experience that has been found by all other members ( $g_{\text{best}}$ ). Therefore, the particles fly around in a multidimensional space towards the better area over the search process.

Each particle consists of three vectors: the position for  $i$ th individual particle can be denoted as  $X_i = (x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(D)})$ , the best previous position  $p_{\text{best}}$  that the  $i$ th particle has searched is  $P_i = (p_i^{(1)}, p_i^{(2)}, \dots, p_i^{(D)})$ , and the fly velocity of the  $i$ th is  $V_i = (v_i^{(1)}, v_i^{(2)}, \dots, v_i^{(D)})$ . The performance of each particle is measured using a fitness

function varying from problem in hand. During the iterative procedure, the  $i$ th particle at iteration  $t$  is updated by

$$\begin{aligned} v_i^d(t+1) &= \omega' \times v_i^d(t) + c_1 \times \varphi_1 \times [p_i^d(t) - x_i^d(t)] \\ &\quad + c_2 \times \varphi_2 \times [p_g^d(t) - x_i^d(t)], \\ x_i^d(t+1) &= x_i^d(t) + v_i^d(t+1), \end{aligned} \quad (5)$$

where  $\omega$  is called inertia weight,  $c_1$  and  $c_2$  are acceleration constants, and  $\varphi_1$  and  $\varphi_2$  are stochastic value of  $[0, 1]$ . In a PSO system, particles change their positions at each time step until a relatively unchanging position has been encountered or a maximum number of iterations have been met. In general, the performance of each particle is measured according to a fitness function, which is problem dependent. In MLR model, (4) is the fitness function under consideration. Figure 1 shows the flowchart of the developed PSO algorithm. For further details regarding PSO, please refer to Kennedy and Eberhart [27] and Bratton and Kennedy [26].

**2.5. Wavelet Analysis.** Wavelet transformations provide useful decomposition of original time series by capturing useful information on various decomposition levels. Discrete wavelet transformation (DWT) is preferred in most of the forecasting problems because of its simplicity and ability to compute with less time. The DWT can be defined as

$$\psi_{m,n}\left(\frac{t-\tau}{s}\right) = \frac{1}{\sqrt{s_0^{m/2}}} \psi\left(\frac{t-n\tau_0 s_0^m}{s_0^m}\right), \quad (6)$$

where  $m$  and  $n$  are integers that control the scale and time. The most common choices for the parameters  $s_0 = 2$  and  $\tau_0 = 1$ .  $\psi(t)$  called the mother wavelet can be defined as  $\int_{-\infty}^{\infty} \psi(t) dt = 0$ .

For a discrete time series  $x(t)$  where  $x(t)$  occurs at discrete time  $t$ , the DWT becomes

$$W_{m,n} = 2^{-m/2} \sum_{t=0}^{N-1} \psi(2^{-m}t - n) x(t), \quad (7)$$

where  $W_{m,n}$  is the wavelet coefficient for the discrete wavelet at scale  $s = 2^m$  and  $\tau = 2^m n$ . According to 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 [28])

$$x(t) = T + \sum_{m=1}^M \sum_{t=0}^{2^{M-m}-1} W_{m,n} 2^{-m/2} \psi(2^{-m}t - n) \quad (8)$$

or in a simple format as

$$x(t) = A_M(t) + \sum_{m=1}^M D_m(t), \quad (9)$$

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

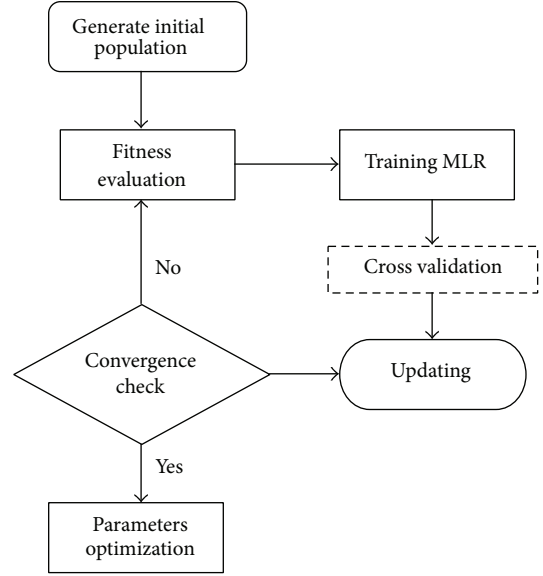


FIGURE 1: Flowchart of PSO algorithm.

**2.6. Principal Component Analysis.** In an MLR, one of main tasks is to determine the model input variables that affect the output variables significantly. The choice of input variables is generally based on a priori knowledge of causal variables, inspections of time series plots, and statistical analysis of potential inputs and outputs. PCA is a technique widely used for reducing the number of input variables when we have huge volume of information and we want to have a better interpretation of variables (Çamdevýren et al. [29]).

The PCA approach introduces a few combinations for model input in comparison with the trial and error process. Given a set of centred input vectors  $x_1, x_2, \dots, x_m$  and  $\sum_{t=1}^m x_t = 0$ , usually  $n < m$ . Then the covariance matrix of vector is given by

$$C = \frac{1}{l} \sum_{t=1}^l x_t x_t^T. \quad (10)$$

The principal components (PCs) are computed by solving the eigenvalue problem of covariance matrix  $C$ ,

$$\lambda_i u_i = C u_i, \quad i = 1, 2, \dots, m, \quad (11)$$

where  $\lambda_i$  is one of the eigenvalues of  $C$  and  $u_i$  is the corresponding eigenvector. Based on the estimated  $u_i$ , the components of  $z_t(i)$  are then calculated as the orthogonal transforms of  $x_t$ :

$$z_t(i) = u_i^T x_t, \quad i = 1, 2, \dots, m. \quad (12)$$

The new components,  $z_i(t)$ , are called principal components. By using only the first several eigenvectors sorted in descending order of the eigenvalues, the number of principal components in  $z_t$  can be reduced. So PCA has the dimensional reduction characteristic. The principal components of PCA have the following properties:  $z_t(i)$  are linear combinations of the original variables, uncorrelated and have sequentially

maximum variances (Jolliffe [30]). The calculation variance contribution rate is

$$V_i = \frac{\lambda_i}{\sum_{i=1}^m \lambda_i} \times 100\%. \quad (13)$$

The cumulative variance contribution rate is

$$V(p) = \sum_{i=1}^p V_i. \quad (14)$$

The number of the selected principal components is based on the cumulative variance contribution rate, which as a rule is over 85~90.

### 3. Computer Simulation

**3.1. An Application.** In this study, the West Texas Intermediate (WTI) crude oil price series was chosen as experimental sample. The main reason of selecting the WTI crude oil is that these crude oil prices are the most famous benchmark prices, which are widely used as the basis of many crude oil price formulae. 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 WMLR 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 2 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, three main performance criteria are used to evaluate the accuracy of the models. These criteria are mean absolute error (MAE), root mean squared error (RMSE), and  $D_{\text{stat}}$ . The MAE and RMSE can be defined by

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}, \quad (15)$$

$$\text{MAE} = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|.$$

In crude oil price forecasting, improved decisions usually depend on correct forecasting of directions, of actual price,  $y_t$  and forecasted price,  $\hat{y}_t$ . The ability to predict movement direction can be measured by a directional statistic ( $D_{\text{stat}}$ ) (Yu et al., [1]), which can be expressed as

$$D_{\text{stat}} = \frac{1}{N} \sum_{t=1}^n a_t \times 100\%, \quad (16)$$

$$a_t = \begin{cases} 1, & \text{if } (y_{t+1} - y_t)(\hat{y}_{t+1} - \hat{y}_t) \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

**3.2. Application and Result.** At first, the MLR model without data preprocessing was used to model daily oil prices. In

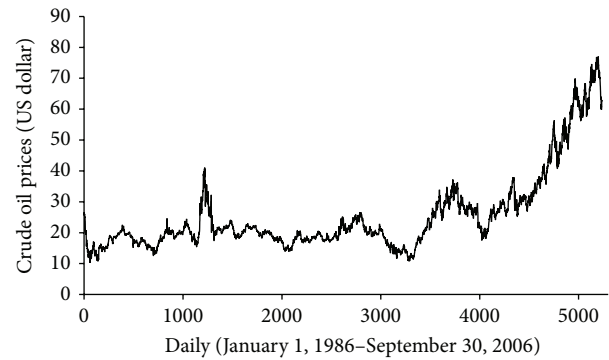


FIGURE 2: Daily crude oil prices from January 1, 198, to September 30, 2006.

the next step, the preprocessed data which uses subtime series components obtained using discrete wavelet transform (DWT) on original data were entered to the MLR model in order to improve the model accuracy. For the MLR model, the original log return time series are decomposed into a certain number of subtime series components. Deciding the optimal decomposition level of the time series data in wavelet analysis plays an important role in preserving the information and reducing the distortion of the datasets. However, there is no existing theory to tell how many decomposition levels are needed for any time series.

In the present study, the previous log return of daily oil price time series is decomposed into various subtime series (DWs) at different decomposition levels by using DWT to estimate current price value. Three decomposition levels (2, 4, and 8 months) were considered for this study. For the WTI series data, time series of 2-day mode (DW1), 4-day mode (DW2) and 8-day mode (DW3), and approximate mode are presented in Figure 3.

For the WTI series, six input combinations based on previous log return of daily oil prices are evaluated to estimate current prices value. The input combinations evaluated in the study are (i)  $r_{t-1}$ , (ii)  $r_{t-1}, r_{t-2}$ , (iii)  $r_{t-1}, r_{t-2}, r_{t-3}$ , (iv)  $r_{t-1}, r_{t-2}, r_{t-3}, r_{t-4}$ , (v)  $r_{t-1}, r_{t-2}, r_{t-3}, r_{t-4}, r_{t-5}$ , and (vi)  $r_{t-1}, r_{t-2}, r_{t-3}, r_{t-4}, r_{t-5}, r_{t-6}$ . In all cases, the output is the log return of current oil prices,  $r_t$ .

Each of DWs series plays distinct role in original time series and has different effects on the original prices oil series. The selection of dominant DWs as inputs of MLR model becomes important and effective on the output data and has positive effect excessively on model's ability. The model becomes exponentially more complex as the number of subtime series as input variables increases. Using a large number of input variables should be avoided to save time and calculation effort. Therefore, the effectiveness of new series obtained by PCA is used as input to the MLR model. The PCA approach helps us to reduce the number of original variables to a set of new variables. Generally, the objective of PCA is to identify a new set of variables such that each variable, called a principal component, is a linear combination of the original variables. The new set of variables accounts for 85%–90% of

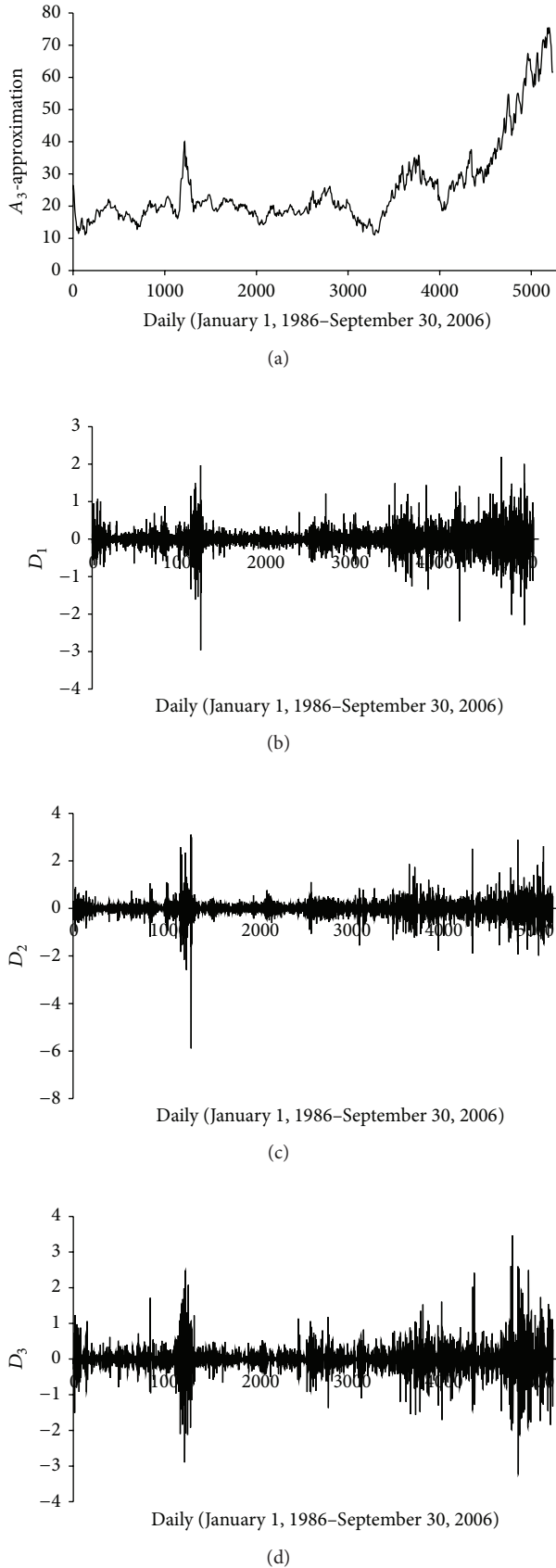


FIGURE 3: Decomposed wavelet subtime series components ( $D_s$ ) of WTI crude oil price data.

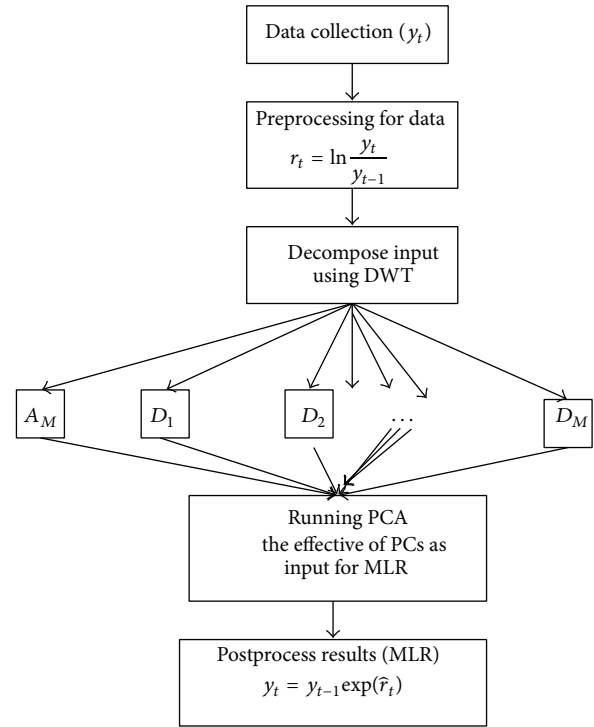


FIGURE 4: The structure of the WMLR model.

TABLE 1: Eigen value and cumulative variance contribution rate of the 8 principal components.

PC	1	2	3	4	5	6	7	8
Eigen value	1.97	1.79	1.59	1.33	0.67	0.41	0.21	0.03
Cumulative Variance Rate	0.25	0.47	0.67	0.84	0.92	0.97	1.00	1.00

the total variation were considered as the number of new variables.

For example, taking two previous daily oil prices as a random variable. Every previous daily oil price time series are decomposed using DWT into three decomposition levels, respectively. Thus there were 8 subseries considered for the PCA analysis. The result of PCA analysis is shown in Table 1. Table 1 shows that the first four principle components can explain 84% variation of the data variation with the eigenvalues greater than 1 to be retained, in which all the 4 PCs were included in the MLR model. Thus the 8 original variables can be replaced by 4 new irrelevant variables. For training MLR, the PSO algorithm solving the recognition problem is implemented and the program code including wavelet toolbox was written in MATLAB language. The WMLR model structure developed in present study is shown in Figure 4.

The forecasting performances of the MLR and WMLR models in terms of the MAE, RMSE, and  $D_{\text{stat}}$  testing phase are compared and shown in Table 2. Table 2 shows MLR model; the M1 with 1 lag obtained the best MAE statistics of 0.6948 and the M6 with 6 lags obtained the best RMSE statistics of 0.9450, while the M1 with 5 lags obtained the best



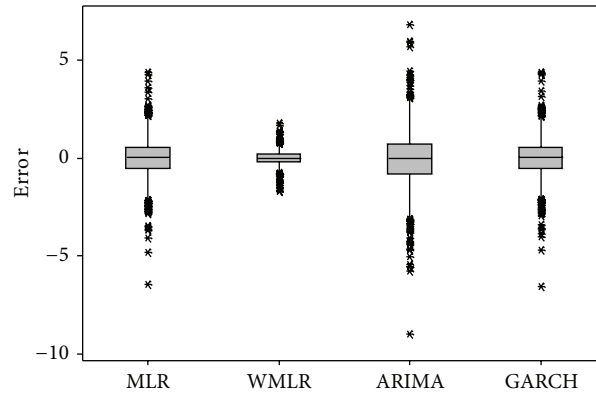


FIGURE 5: The errors of MLR, WMLR, ARIMA, and GARCH models for crude oil price forecasting.

TABLE 2: Forecasting performance indices of MLR and WLR.

Model Input	Lag	MLR			WMLR		
		MAE	RMSE	$D_{stat}$	MAE	RMSE	$D_{stat}$
M1	1	<b>0.6948</b>	0.9514	0.4788	0.6660	0.9001	0.5198
M2	1, 2	0.6972	0.9517	0.4781	0.6448	0.8842	0.5003
M3	1, 2, 3	0.6985	0.9545	0.4816	0.5345	0.7505	0.5797
M4	1, 2, 3, 4	0.6979	0.9550	0.4753	<b>0.4834</b>	<b>0.6572</b>	<b>0.6722</b>
M5	1, 2, 3, 4, 5	0.6976	0.9545	<b>0.4878</b>	0.5770	0.8046	0.5734
M6	1, 2, 3, 4, 5, 6	0.6969	<b>0.9450</b>	0.4850	0.5385	0.7389	0.6444

$D_{stat}$  statistics of 0.4878. For WMLR, model M4 with 4 lags obtained the best MAE, RMSE, and  $D_{stat}$  statistics of 0.4834, 0.6572, and 0.6722, respectively. The equations of MLR with six input variables and WMLR with four input variables, respectively, are

$$\begin{aligned}
 \hat{y}_t = & -0.025y_{t-1} - 0.047y_{t-2} + 0.22y_{t-3} \\
 & - 0.082y_{t-4} - 0.060y_{t-5} + 0.0004y_{t-6}, \\
 r_t = & 0.076z_1(t) - 0.221z_2(t) - 0.213z_3(t) \\
 & - 0.510z_4(t) + 0.416z_5(t) - 0.930z_6(t),
 \end{aligned} \tag{17}$$

where  $z_i(t)$  are called principal components and  $\hat{y}_t = y_{t-1} \exp(r_t)$ .

For further analysis, the best performance of the LR, WMLR, ARIMA, and ARIMA-GARCH models was compared with the best results of ARIMA and forward neural network (FNN) studied by Yu et al. [1]. In Table 3, it shows that WMLR outperform MLR, ARIMA, GARCH, Yu' ARIMA and Yu' FNN models in terms of RMSE statistics. This results show that the new series (DWT) have significant extremely positive effect on MLR model results.

Figure 5 shows the Box-plot for the ARIMA, ARIMA-GARCH, MLR, and WMLR models for testing period. It can be seen that the errors of WMLR model are quite close to the zero. Overall, it can be concluded that the WMLR model provided more accurate forecasting results than the other models for crude oil forecasting.

TABLE 3: The RMSE and MAE comparisons for different models.

Model	RMSE	MAE
ARIMA (2, 1, 5)	1.3835	1.0207
GARCH (1, 1)	0.9513	0.6947
MLR	0.9450	0.6969
WMLR	<b>0.6572</b>	<b>0.4834</b>
Yu' ARIMA (Yu et al., [1])	2.0350	—
Yu' FNN (Yu et al., [1])	0.8410	—

## 4. Conclusions

The accuracy of the wavelet multiple linear regression (WMLR) technique in the forecasting daily crude oil has been investigated in this study. The PCA is used to choose the principle component scores of the selected inputs which were used as independent variables in the MLR model and the particle swarm optimization (PSO) is used to adopt the optimal parameters of the MLR model. The performance of the proposed WMLR model was compared to regular LR, ARIMA, and GARCH model for crude oil forecasting. Comparison results indicated that the WMLR model was substantially more accurate than the other models. The study concludes that the forecasting abilities of the MLR model are found to be improved when the wavelet transformation technique is adopted for the data preprocessing. 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 MLR model. The accurate forecasting results

indicate that WMLR model provides a superior alternative to other models and a potentially very useful new method for crude oil forecasting. The WMLR model presented in this study is a simple explicit mathematical formulation. The WMLR model is much simpler in contrast to ANN model and can be successfully used in modeling short-term crude oil price. In the present study, three resolution levels were employed for decomposing crude oil time series. If more resolution levels were used, the results from WMLR model may turn out better. This may be a subject of another study.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

The authors thankfully acknowledged the financial support that was afforded by Universiti Teknologi Malaysia under GUP Grant (VOT 06J13). Besides that, the authors would like to thank the Department of Irrigation and Drainage, Ministry of natural Resources and Environment, Malaysia, in helping us to provide the data.

## References

- [1] L. Yu, S. Wang, and K. K. Lai, "Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm," *Energy Economics*, vol. 30, no. 5, pp. 2623–2635, 2008.
- [2] H. Mohammadi and L. Su, "International evidence on crude oil price dynamics: applications of ARIMA-GARCH models," *Energy Economics*, vol. 32, no. 5, pp. 1001–1008, 2010.
- [3] M. I. Ahmad, "Modelling and forecasting Oman Crude Oil Prices using Box-Jenkins techniques," *International Journal of Trade and Global Markets*, vol. 5, pp. 24–30, 2012.
- [4] P. Agnolucci, "Volatility in crude oil futures: a comparison of the predictive ability of GARCH and implied volatility models," *Energy Economics*, vol. 31, no. 2, pp. 316–321, 2009.
- [5] Y. Wei, Y. Wang, and D. Huang, "Forecasting crude oil market volatility: further evidence using GARCH-class models," *Energy Economics*, vol. 32, no. 6, pp. 1477–1484, 2010.
- [6] L. Liu and J. Wan, "A Study of Shanghai fuel oil futures price volatility based on high frequency data: long-range dependence, modeling and forecasting," *Economic Modelling*, vol. 29, pp. 2245–2253, 2012.
- [7] A. Ghaffari and S. Zare, "A novel algorithm for prediction of crude oil price variation based on soft computing," *Energy Economics*, vol. 31, no. 4, pp. 531–536, 2009.
- [8] M. A. Kaboudan, "Computetric forecasting of crude oil prices," in *The Proceedings of IEEE Congress on Evolutionary Computation*, pp. 283–287.
- [9] S. Mirmirani and H. C. Li, "A comparison of VAR and neural networks with genetic algorithm in forecasting price of oil," *Advances in Econometrics*, vol. 19, pp. 203–223, 2004.
- [10] W. E. Shambora and R. Rossiter, "Are there exploitable inefficiencies in the futures market for oil?" *Energy Economics*, vol. 29, no. 1, pp. 18–27, 2007.
- [11] L. Yu, S. Wang, and K. K. Lai, "A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates," *Computers and Operations Research*, vol. 32, no. 10, pp. 2523–2541, 2005.
- [12] W. Xie, L. Yu, S. Y. Xu, and S. Y. Wang, "A new method for crude oil price forecasting based on support vector machines," *Lecture Notes in Computer Science*, vol. 3994, pp. 441–451, 2006.
- [13] R. Jammazi and C. Aloui, "Crude oil price forecasting: experimental evidence from wavelet decomposition and neural network modeling," *Energy Economics*, vol. 34, no. 3, pp. 828–841, 2012.
- [14] W. Qunli, H. Ge, and C. Xiaodong, "Crude oil price forecasting with an improved model based on wavelet transform and RBF neural network," in *Proceedings of the International Forum on Information Technology and Applications (IFITA '09)*, pp. 231–234, May 2009.
- [15] S. Yousefi, I. Weinreich, and D. Reinartz, "Wavelet-based prediction of oil prices," *Chaos, Solitons and Fractals*, vol. 25, no. 2, pp. 265–275, 2005.
- [16] Y. Bao, X. Zhang, L. Yu, K. K. Lai, and S. Wang, "Hybridizing wavelet and least squares support vector machines for crude oil price forecasting," in *Proceedings of the 2nd International Workshop on Intelligent Finance*, Chengdu, China, 2007.
- [17] J. Liu, Y. Bai, and B. Li, "A new approach to forecast crude oil price based on fuzzy neural network," in *Proceedings of the 4th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD '07)*, pp. 273–277, August 2007.
- [18] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, Calif, USA, 1976.
- [19] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, vol. 31, no. 3, pp. 307–327, 1986.
- [20] H. Bozdogan and J. A. Howe, "Misspecified multivariate regression models using the genetic algorithm and information complexity as the fitness function," *European Journal of Pure and Applied Mathematics*, vol. 5, no. 2, pp. 211–249, 2012.
- [21] S. M. Chen and P. Y. Kao, "TAEIX forecasting based on fuzzy time series, partial swarm optimization techniques and support vector machines," *Information Sciences*, vol. 247, pp. 62–71, 2013.
- [22] R. Alwee, S. M. Shamsuddin, and R. Sallehuddin, "Hybrid support vector regression and autoregressive integrated moving average models improved by particle swarm optimization for property crime rates forecasting with economic indicators," *The Scientific World Journal*, vol. 2013, Article ID 951475, 11 pages, 2013.
- [23] R. J. Ma, N. Y. Yu, and J. Y. Hu, "Application of particle swarm optimization algorithm in the heating system planning problem," *The Scientific World Journal*, vol. 2013, Article ID 718345, 11 pages, 2013.
- [24] G. Tyagi and M. Pandit, "Combined heat and power dispatch using time varying acceleration coefficient particle swarm optimization," *International Journal of Engineering and Innovative Technology*, vol. 1, no. 4, pp. 234–240, 2012.
- [25] L. Liu, S. Yang, and D. Wang, "Particle swarm optimization with composite particles in dynamic environments," *IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics*, vol. 40, no. 6, pp. 1634–1648, 2010.
- [26] D. Bratton and J. Kennedy, "Defining a standard for particle swarm optimization," in *Proceedings of the IEEE Swarm Intelligence Symposium (SIS '07)*, pp. 120–127, IEEE Service Center, Piscataway, NJ, USA, April 2007.

- [27] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Conference on Neural Networks*, pp. 1942–1948, December 1995.
- [28] S. G. Mallat, "Theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.
- [29] H. Çamdevýren, N. Demýr, A. Kanik, and S. Keskýn, "Use of principal component scores in multiple linear regression models for prediction of Chlorophyll-a in reservoirs," *Ecological Modelling*, vol. 181, no. 4, pp. 581–589, 2005.
- [30] I. T. Jolliffe, *Principal Component Analysis*, Springer, New York, NY, USA, 1986.



