Time Series Forecasting of Energy Commodity using Grey Wolf Optimizer

Yuhanis Yusof, *Member, IAENG* and Zuriani Mustaffa

***Abstract*— The ability to model and perform decision making is an essential feature of many real-world applications including the forecasting of commodity prices. In this study, a forecasting model based on a relatively new Swarm Intelligence (SI) behaviour, namely Grey Wolf Optimizer (GWO), is developed for short term time series forecasting. The model is built upon data obtained from the West Texas Intermediate (WTI) crude oil and gasoline price. Performance of the GWO model is compared against two other models which are developed based on Evolutionary Computation (EC) algorithms, namely the Artificial Bee Colony (ABC) and Differential Evolution (DE). Results showed that the GWO model outperformed DE in both crude oil and gasoline price forecasting. Furthermore, the proposed GWO produces a better forecast for gasoline price as compared to the ABC model,, as well as being at par in crude oil. Such an achievement indicates that GWO may become a competitor in the domain of time series forecasting and would be useful for investors in planning their investment and projecting their profit.**

***Index Terms*— time series forecasting, Grey Wolf Optimizer, Artificial Bee Colony, swarm intelligence, data mining**

I. Introduction

Forecasting crude oil price is proven to be challenging and of great interest to practitioners, governments, enterprises and academia. Known as 'black gold' due to its prosperous characteristics, it is regarded as one of the most significant resources as it has the strength to influence world economic development [1]. A reliable forecasting tool for the said time series data is not only essential in avoiding unwanted risk, reducing loss and gaining high profit but also contributes to an appropriate future planning. Possible development to overcome expected issue can be taken into account. Nonetheless, due to high complexity and nonlinearity features which caused by various factors such as supply and demand inventory, political situation, inflation, Gross Domestic Product (GDP) and many others, the price is continued to be hard to forecast [2, 3].

Manuscript received January 18, 2015; revised January 21, 2015. This work was supported by the Malaysian Ministry of Education under Grant ERGS s/o code (12467).

Yuhanis Yusof is with the School of Computing, College of Arts and Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah, MALAYSIA; phone: +60133921224; fax: +6049285067; e-mail: yuhanis@uum.edu.my.

Zuriani Mustaffa was with School of Computing, Universiti Utara Malaysia. . She is now with the Faculty of Computer Systems & Software Engineering, Universiti Malaysia Pahang, Malaysia (e-mail: zuriani@ump.edu.my).

Classified as non renewable natural resources commodity, crude oil is very limited in production and irreplaceable in human time frame [4]. With the limitation in resources and continuously increasing demand, this situation leads to only one result; higher prices. As for investors, this means opportunity, however, for public people, this indicates inflation [1, 5]. Due to that matter, the importance of price forecasting for such data has resulted to a large growing body of literature and research among the community is continuously carried out [3].

In literature, there are avalanche of studies which present various forecasting techniques for the said time series data. In [3], monthly crude oil price forecasting was implemented based on an improved Back Propagation Neural Network (BPNN). Realized in West Texas Intermediate (WTI) crude oil price, the BPNN model is compared against conventional BPNN. The finding of the study was in favour to the improved BPNN. Meanwhile, a hybridization of Genetic Algorithm and Feed Forward Neural Network (FFNN) with BP algorithm has also been demonstrated in crude oil price forecasting [6]. In the study, GA was employed to improve the learning algorithm and reduce the complexity in determining the control parameters of ANN. Later, the prediction process is continued by the FFNN. The experimental process involved two time series data of crude oil prices, viz. WTI and Iran crude oil prices and comparison was conducted against conventional Artificial Neural Network (ANN). Upon completing the experiment, it is indicated that the results produced by GA-FFNN are closer to actual data.

Progressing further, an ensemble machine learning technique was evaluated in forecasting crude oil price [7, 8]. In the study, three machine learning algorithms were chosen for comparison purposes which include Support Vector Machine (SVM), Instant Based Learning (IBL) and K.Star. Empirical results suggested that the developed ensemble algorithm performed better than the identified forecasting algorithm. In related work, the combination of Pattern Modelling and Recognition System (PMRS), Error Correction model (ECM) and Neural Networks (NN) has been presented to forecast the monthly WTI crude oil price [9]. The empirical results suggested that the presented model give good forecasting performance relative to the Mean Absolute Percentage Error (MAPE) and Root Mean Square Percentage Error (RMSPE). These methods, to a certain extent, all improve the accuracy of crude oil price forecasting.

In [10] the researchers attempt to predict crude oil price using Empirical Mode Decomposition (EMD)-based Neural Network ensemble learning paradigm. In the study, the prediction task is done by using Adaptive Linear Neural Network(ALNN). Firstly, the crude oil price was first decomposed into a number of Intrinsic Mode Function (IMF). Later, the ALNN is used to predict each of the IMF. For each prediction results, a weight is assigned and then, the obtained results are combined together. The evaluation is later made based on RMSE and Dstat. Even empirical results of the study showed an encouraging result, however, this approach is ineffective since if one of the ALNN yields a poor prediction results, the results might be affected since all results are summed together. This may result in imprecise prediction [11].

Nonetheless, despite the various presented techniques in crude oil price forecasting, finding an effective forecasting model for the said time series data is important. The gaps that exist in existing work, particularly the Neural Network based model [3, 6, 12] which is favorably applied in crude oil price forecasting is unavoidable to face with the poor generalization [13, 14] and the requirement of many control parameters to be tuned [14, 15]. In this study, Grey Wolf Optimizer (GWO) [16] is developed to forecast daily crude oil prices. As a relatively new Swarm Intelligence (SI) algorithm, GWO is motivated from social behaviour of grey wolves or also known as Canis Lupus which belongs to Canidae group. This algorithm consists of four main parts namely social hierarchy, encircling prey, hunting, attacking prey and search for prey. Similarly like any other meta heuristic algorithm, exploitation and exploration are also the two important features of GWO. In GWO, these features are reflected in the attacking prey and search for prey respectively. To date, GWO algorithm has been proven to be competitive and better than the other existing optimization algorithms such as Particle Swarm Optimization (PSO), Gravitational Search Algorithm (GSA) and many others [16]. With such performance, the GWO posses a great potential for forecasting the non renewable time series data. This crude oil time series data is chosen due to its significant role not only in human life survival but also contributes to the global economic activities. In forecasting, GWO is used to identify optimal values of the parameters in the prediction function, as applied in existing work [17, 18].

II. GREY WOLF OPTIMIZER (GWO)

A. Theory of GWO

GWO is considered as apex predators, which makes them placed as the top in food chain. In GWO, there are 4 hierarchies in grey wolf population, namely alpha, beta delta and omega. In the alpha level, it consists of male and female grey wolf and is responsible for decision making on hunting, sleeping place and others. Due to its dominant role, they are placed at the top of the hierarchy. The second level, beta, is responsible to help the alpha in decision making or any other activities of the pack. The beta can be male or female and will be the best candidate in replacing the alpha if one of the alpha passes away or become old. The beta acts as an advisor for the alpha in undertaking discipline of the pack. Meanwhile, wolf placed at the delta level are required to forward solutions to alpha and beta but they dominate the omega. This group consist of scouts, sentinels, elders, hunters and caretakers. Lastly, the omega, which is ranked last in the hierarchy, plays the role as scapegoat.

B. Mathematical Model and Algorithm

Social Hierarchy

In GWO, the fittest solution is represented by alpha (α), followed by the second and third best solutions which are the beta (β) and delta (δ) respectively. Meanwhile, the balance of the candidate solutions is considered as omega (ω). The hunting (optimization) is guided by α, β and δ while the ω follows the three previous groups.

Encircling Prey

During hunting, the wolves tend to encircle their prey. As to model the encircling prey, the following equation is used:

$$\vec{D} = |\vec{C}.\vec{X}_p(t) - \vec{X}(t)| \quad (1)$$

$$\vec{X}(t+1) = \vec{X}_p(t) - \vec{A}.\vec{D} \quad (2)$$

where t *=current iteration,* $\vec{A}$ *and* $\vec{C}$ *= coefficient vectors,* $\vec{X}_p$ *= position vector of the prey and* $\vec{X}$ *= position vector of the grey wolves.*

For vectors $\vec{A}$ and $\vec{C}$, it is calculated as follows:

$$\vec{A} = 2\vec{a}.\vec{r}_1 - \vec{a} \quad (3)$$

$$\vec{C} = 2.\vec{r}_2 \quad (4)$$

where components of $\vec{a}$ are linearly decreased from 2 to 0 over the curse of iterations. Meanwhile, r_1 and r_2 are random vectors in the range of [0,1].

Hunting

Commonly, the hunting is guided by the alpha. However, both beta and delta may also be involved in hunting, occasionally. In GWO, the alpha, i.e. the fittest candidate solution, beta and delta are the experts about the potential location of prey. Thus, the first three best solutions obtained are stored while the other agents (including omegas) are induced to update their positions based on the position of the best search agents. This is defined by:

$$\vec{D}_\alpha = |\vec{C}_1.\vec{X}_\alpha - \vec{X}|, \vec{D}_\beta = |\vec{C}_2.\vec{X}_\beta - \vec{X}|, \vec{D}_\delta = |\vec{C}_3.\vec{X}_\delta - \vec{X}| \quad (5)$$

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1.(\vec{D}_\alpha), \vec{X}_2 = \vec{X}_\beta - \vec{A}_2.(\vec{D}_\beta), \vec{X}_3 = \vec{X}_\delta - \vec{A}_3.(\vec{D}_\delta) \quad (6)$$

ISBN: 978-988-19253-2-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \quad (7)$$

Details on the GWO theory can be seen in [16]..

III. METHODOLOGY

This section elaborates steps taken in developing GWO forecasting models; crude oil price forecasting and gasoline price forecasting. Upon completing the data collection and pre-process stage, the forecasting algorithms were designed and developed. Evaluation of the forecasting models was then undertaken by comparing their results against the ones produced by state of the art in forecasting models.

A. Research Data and Data Preparation

In this study, real data of West Texas Intermediate (WTI) crude oil and gasoline prices are utilized in the experiments. Such datasets are included as they are the benchmark datasets in price forecasting [5]. The time series data covered in this experiment starts from December 1, 1997 to June 30, 1998 and is obtained from Barchart website [19]. Table 1 includes examples of raw data obtained from the website.

Table 1. Sample of Raw Data

Date	CL Price	HU Price
12/1/1997	18.6300	0.5338
12/2/1997	18.7000	0.5316
12/3/1997	18.6000	0.5328
12/4/1997	18.5900	0.5272
12/5/1997	18.7000	0.5262

Prior to dividing the dataset into training and testing sets, the data in-hand is pre-processed. This is done by deriving additional statistical attributes as undertaken in [16, 20] that includes percentage of change in the commodity price from the previous day, standard deviation over the previous 5 and 21 working days of the commodity price. From the sample, 70% of the dataset is allocated for training purposes while the balance of 30% is utilized as testing.

The variables assigned to features involved in predicting crude oil and gasoline price are as tabulated in Table 2 and Table 3. The undertaken experiment utilizes the daily spot price of crude oil for one month ahead (21 trading days) as the output.

Table 2. Input and Output Variables for Crude Oil Model

Input	Variable	Output
Daily closing price of crude oil	CL	Daily spot price of crude oil from day 21 onwards (CL21)
Percent change in crude oil daily closing spot price from the previous day	%Chg	
Standard deviation over the previous 5 trading days of crude oil price	Std5	
Standard deviation over the previous 21 trading days of crude oil price	Std21	

Table 3. Input and Output Variables for Gasoline Model

Input	Variable	Output
Daily closing price of gasoline	HU	Daily spot price of gasoline from day 21 onwards (HU21)
Percent change in gasoline daily closing spot price from the previous day	%Chg	
Standard deviation over the previous 5 trading days of gasoline price	Std5	
Standard deviation over the previous 21 trading days of gasoline price	Std21	

B. GWO for Price Forecasting

In this forecasting study, the goal is to minimize the error between the forecast and actual price of the energy commodity (i.e crude oil or gasoline). For that purpose, the objective function is served by Mean Absolute Percentage Error (MAPE). The equation for crude oil price forecasting is adapted from [17] and is defined as equation 8:

$$CL21 = (\alpha \times CL) + (\beta \times \%Chg) + (\gamma \times Std5) + (\delta \times Std21) + \varepsilon \quad (8)$$

where the α, β, γ and δ are the coefficients for CL, %Chg, Std5 and Std21 respectively (see Table 1) while the ε is the intercept coefficient.

On the other hand, equation 9 depicts the relevant function for gasoline price forecasting.

$$HU21 = (\alpha \times HU) + (\beta \times \%Chg) + (\gamma \times Std5) + (\delta \times Std21) + \varepsilon \quad (9)$$

where the α, β, γ and δ are the coefficients for HU, %Chg, Std5 and Std21 respectively (see Table 2) while the ε is the intercept coefficient.

The GWO algorithm in forecasting is given in Algorithm 1.

Algorithm **1** GWO algorithm

1: Initialize the population
2: Initialize a, A and C
3: Evaluate the fitness value of parameters of interest using equation (8) or (9)
4: **while** Cycle <= MCN
5: **for** each search agent
6: Update the position of the current search agent using equation(7)
7: **end for**
8: update a, A and C
9: Evaluate parameters of interest and calculate the fitness value using equation (8) or (9)
9: Update X_α, X_β and X_δ
10:Cycle = Cycle + 1
11:**end while**
12:return X_a
13:Print optimal parameters
14:Obtain prediction results

ISBN: 978-988-19253-2-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

C. *Evaluation of GWO in Price Forecasting*

In this study, results from the GWO forecasting model are compared with the results produced by state of the art models; Artificial Bee Colony (ABC) and Differential Evolution (DE) algorithms. Differential Evolution algorithm was introduced by Storn and Price (1997) and is inspired by the mechanism of natural selection which is considered as an extension of Genetic Algorithm (GA). The difference between DE and GA is that in the first algorithm, all possible solutions have an equal chance in the evaluation task, while in the latter algorithm; the chance of updating a solution relies on the fitness value. The adapted DE for forecasting is included in Algorithm 2.

Algorithm 2: DE algorithm

1: Initialize possible solutions
2: Evaluate using equation (8) or (9)
3: Set the weight F and crossover probability
4: while Cycle <= *MCN*
5: for i =1 to n
6: for each x_i, randomly choose 3 distinct vectors x_p, x_r and x_r
7: Generate a new vector v
8: Generate a random index J_r by permutation
9:Generate a randomly distributed number r_i
10:for j = to d
11:for each parameter $v_{j,i}$ (jth component of v_i), update

$$u_{j,i}^{t+1} = \begin{cases} v_{j,i}^{t+1} & \text{if } r_i \leq C_r \text{ or } j = J_r \\ x_{j,i}^{t} & \text{if } r_i > C_r \text{ and } j \neq J_r \end{cases}$$

12:Evaluate using equation (8) or (9)
13:end
14:Select and update the solution
15:Evaluate using equation (8) or (9)
16:end
17:Update the counters
18:end
19:Print optimal parameters
20:Obtain results

On the other hand, ABC algorithm which has been introduced by Dervis Karaboga [21] is enlightened from the intelligent foraging behaviour of honey bees swarm. The ABC consists of three groups of bees viz. employed bee, onlooker bees and scout bees. Half of the colony is omposed of the EB and the rest are filled with the OB. Meanwhile, the SB is basically an EB which change the status for certain condition, which is described later. The number of food sources/nectar sources is equal to the EB. This means that one EB is associated for a single nectar source. The goal of the whole colony is to maximize the amount of nectar.In realizing ABC in forecasting, this study refers to the one presented in Algorithm 2 [22].

The performance of the forecasting models is then evaluated via statistical evaluation indices; Mean Absolute Percentage Error (MAPE) [23] and prediction accuracy (PA). Definitions of these evaluation metrics are shown as follows:

$$MAPE = \frac{1}{N}\left[\sum_{n=1}^{N}\left|\frac{y_n - y(x_n)}{y_n}\right|\right] \tag{10}$$

$$PA = 100\% - (MAPE \times 100) \tag{11}$$

Algorithm 3 ABC algorithm

1: Initialize possible solutions
2: Evaluate the fitness value of parameters of interest using equation (8) or (9)
3: Cycle = 1
4: **while** Cycle <= *MCN*
5: **for** each EB **(EB PHASE)**
6: Produce new solution
7: **if** the new solution is out of boundary, shift the new solution value to the boundary
8: Evaluate parameters of interest and calculate the fitness value using equation (8) or (9)
9: **if** fitness value of new solution is better than fitness value of old solution
10:Keep new solution
11:Trial = 0.
12:**else**
13:Keep old solution
14:Trial = Trial +1
15:**end if**
16:**end for**
17:Calculate the probability values for solution
18:**for** each OB**(OB PHASE)**
19:Select a solution based on probability value
20:Modify selected solution
21:**if** the new solution is out of boundary, shift the new solution value to the boundary
22:Evaluate parameters of interest and calculate the fitness value using equation (8) or (9)
23:**if** fitness value of new solution is better than fitness value of old solution
24:Keep new solution
25:Trial = 0
26:**else**
27:Keep old solution
28:Trial = Trial + 1
29:**end if**
30:**end for**
31:**if** **(**max**)** Trial>*Limit*
32 **SB PHASE**
33: Assign responsible EB as SB and produce new solution to replace the abandoned food source
34: Evaluate parameters of interest and calculate the fitness value using equation (8) or (9)
35:**else**
36:Memorized best solution
37:**end if**
38:**end for**
39:Cycle = Cycle +1
40:**end while**
41:Print optimal parameters
42:Obtain results

ISBN: 978-988-19253-2-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

IV. RESULTS AND DISCUSSIONS

For comparison purposes, the forecasting performance of GWO is compared against the results produced by ABC and DE. According to the results depicted in Table 4, the values of α, β γ, δ and ε identified by GWO are 0.8346, 0.1128, 0.1521, 1 and 1 respectively. The combination of these parameters produced a small value of MAPE which is 5.48%, hence, achieving 94.52% accuracy (PA).

Table 4: Results for Crude Oil Price Forecasting

	GWO	ABC	DE
α	0.8346	0.8454	0.8454
β	0.1128	0.1081	0.1081
γ	0.1521	0.4255	0.4255
δ	1.0000	0.8673	0.8673
ε	1.0000	0.9119	0.9119
MAPE Testing(%)	5.4779	5.4170	11.9320
PA(%)	94.5221	94.5830	88.0680

Similarly, the ABC model also generates a MAPE value that is less than 5.5%. Nevertheless, a statistical paired sample T-test showed that the difference between the average value of MAPE produced by GWO and ABC is not significant at 0.05% significance level (refer to Table 5). By obtaining high correlation, which is 0.9902, it indicates that the prediction values produced by both techniques move very much in the same pattern. On the other hand, results produced by the DE model reside at a lower level. The obtained error rate (i.e MAPE) is larger than 10%, which later produce 88% accuracy.

Table 5: Significant Test for Crude Oil Price Forecasting

	Pearson Correlation	Sig. (2-tailed)
GWO - ABC	0.9902	0.6111
GWO - DE	0.3049	0.0003
ABC - DE	0.2913	0.0003

The performance of the models in forecasting crude oil price is also illustrated in Figure 1. The figure plots the actual and forecast value of GWO and the identified competitors from day 103 to day 146 (testing phase). The dashed line represents actual price while the GWO forecast price is indicated by a solid line. On the other hand, the diamond mark and cross mark represent the forecast value obtained using ABC and DE respectively.

The result produced by the second GWO forecasting model is presented in Table 6. In predicting the gasoline price of the same time period, it is learned that the GWO produced a better result. Comparing the three models, GWO obtain the highest accuracy which is 93.15% while ABC has the least value. The undertaken T-test (as shown in Table 7) also reveals that the difference between the mean of predicted price values between GWO and ABC, GWO and DE, and ABC-DE is significant at 0.05 level of significance.

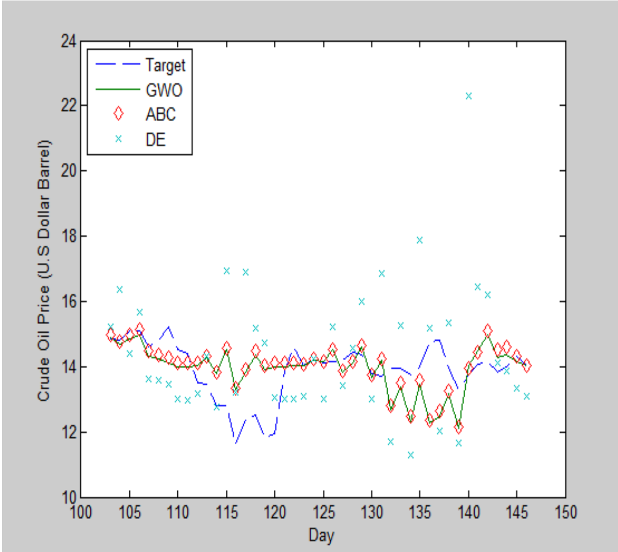

Figure 1: Actual vs. Forecast Values by GWO, ABC and DE

Table 6: Results for Gasoline Price Forecasting

	GWO	ABC	DE
α	0.2312	0.0971	0.3903
β	0.0016	0	0.0017
γ	0	0.3032	0.5266
δ	0.11988	0	0.586
ε	0.3508	0.4252	0.2696
MAPE Testing(%)	6.8481	9.0652	8.0137
PA(%)	93.1519	90.9348	91.9863

Table 7: Significant Test for Gasoline Price Forecasting

	Pearson Correlation	Sig. (2-tailed)
GWO - ABC	0.9802	0.0000
GWO - DE	0.9759	0.0000
ABC - DE	0.9545	0.0004

V. CONCLUSION

Prediction of non-renewable natural commodity has experienced major changes for past decades. Starting from conventional statistical techniques to artificial intelligence approach, this issue has never failed to attract both academic and practitioners community. In this study, a new SI algorithm namely Grey Wolf Optimizer is employed for short term crude oil and gasoline price forecasting. The efficiency of the developed GWO forecasting models is measured based on Mean Absolute Percentage Error and prediction accuracy and is compared against the ones produced by Artificial Bee Colony and Differential Equation models. Findings of the study reveal competitive results where it is learned that the GWO is comparable to ABC algorithm in predicting gold price while becoming a better predictor for gasoline price. Such forecasting model would benefit the investors in planning their investment in energy commodity. As in future, it would be interesting to test the efficiency and applicability of the GWO on renewable commodities such as currencies and stocks.

REFERENCES

[1] Y.J. Zhang and Y.-M. Wei, "The Crude Oil Market and the Gold Market: Evidence for Cointegration, Causality and Price Discovery," *Resources Policy*, vol. 35, pp. 168-177, 2010.

[2] Y. Li and W. Ma, "Applications of Artificial Neural Networks in Financial Economics: A Survey," in *Proceedings of the International Symposium on Computational Intelligence and Design*, 2010.

[3] R. Jammazi and C. Aloui, "Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modeling," *Energy Economics*, vol. 34, pp. 828-841, 2012.

[4] D. D. Kemp, *Exploring Environmental Issues: An Integrated Approach* London: Routledge, 2004.

[5] K. He, L. Yu, and K. K. Lai, "Crude Oil Price Analysis and Forecasting using Wavelet Decomposed Ensemble Model," *Energy Economics*, vol. 46, pp. 564-574, 2012.

[6] R. Tehrani and F. Khodayar, "A hybrid optimized Artificial Intelligent Model to Forecast Crude Oil using Genetic Algorithm," *African Journal of Bussiness Management*, vol. 5, pp. 13130-13135, 2011.

[7] L. A. Gabralla, R. Jammazi, and A. Abraham, "Oil Price Prediction Using Ensemble Machine Learning," in *Proceedings of the International Conference on Computing, Electrical and Electronics Engineering (ICCEEE)*, Khartoum, 2013, pp. 674-679.

[8] M. Espinoza, J. Suykens, and B. Moor, "Fixed-size Least Squares Support Vector Machines: A Large Scale Application in Electrical Load Forecasting," *Computational Management Science*, vol. 3, pp. 113-129, 2006.

[9] D. Xu, Y. Zhang, C. Cheng, W. Xu, and L. Zhang, "A Neural Network-based Ensemble Prediction using PMRS and ECM," in *Proceedings of the 47th Hawaii International Conference on System Science*, Hawaii, 2014.

[10] 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, pp. 2623-2635, 2008.

[11] S. Zhou, K. K. Lai, and J. Yen, "A Dynamic Meta-Learning Rate-Based Model for Gold Market Forecasting," *Expert Systems with Applications*, vol. 39, pp. 6168-6173, 2012.

[12] W. Jian-Zhou, W. Ju-Jie, Z. Zhe-George, and G. Shu-Po, "Forecasting Stock Indices with Back Propagation Neural Network," *Expert Systems with Applications*, vol. 38, pp. 14346-14355, 2011.

[13] J. Cheng, J.-S. Qian, and Y.-N. Guo, "Least Squares Support Vector Machines for Gas Concentration Forecasting in Coal Mine," *International Journal of Computer Science and Network Security*, vol. 6, pp. 125-129, 2006.

[14] Y. Xiang and L. Jiang, "Water Quality Prediction Using LS-SVM and Particle Swarm Optimization," in *Proceedings of the Second International Workshop on Knowledge Discovery and Data Mining (WKDD)* 2009, pp. 900-904.

[15] G. Zhang, B. E. Patuwo, and M. Y. Hu, "Forecasting with Artificial Neural Network: The State of Art," *International Journal of Forecasting*, vol. 14, pp. 35-62, 1998.

[16] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Advance in Engineering Software*, vol. 69, pp. 46-61, 2014.

[17] E. Hadavandi, A. Ghanbari, and S. Abbasian-Naghneh, "Developing a Time Series Model Based on Particle Swarm Optimization for Gold Price Forecasting," in *Proceedings of the Third International Conference on Business Intelligence and Financial Engineering (BIFE)*, 2010, pp. 337-340.

[18] Z. Mustaffa, Y. Yusof, and S. S. Kamaruddin, "Enhanced ABC-LSSVM for Energy Fuel Price Prediction," *Journal of Information and Communication Technology.*, 2013.

[19] (2012, November, 2011). *Barchart*. Available: http://www.barchart.com/historicalquote.php?view=quote&sym=CLY00&txtDate=12%2F31%2F09&submit=Get+Data

[20] M. E. Malliaris and S. G. Malliaris, "Forecasting Inter-Related Energy Product Prices," *The European Journal of Finance*, vol. 14, pp. 453-468, 2008.

[21] D. Karaboga, "An Idea Based on Honey Bee Swarm for Numerical Optimization," Erciyes University, Technical ReportOctober, 2005 2005.

[22] Z. Mustaffa, Y. Yusof, and S. S. Kamaruddin, "Gold Price Prediction using Artificial Bee Colony Algorithm," in *Proceedings of the 2nd International Conference on Engineering and Applied Science (ICEAS)*, Tokyo, Japan, 2013.

[23] R. J. Hyndman and A. B. Koehler, "Another Look at Measures of Forecast Accuracy," *International Journal of Forecasting*, vol. 22, pp. 679-688, 2006.