Application of Grey Wolf Optimizer for Time Series Forecasting

Zuriani Mustaffa, Universiti Malaysia Pahang, Malaysia
Yuhanis Yusof, Universiti Utara Malaysia, Malaysia
Siti Sakira Kamaruddin, Universiti Utara Malaysia, Malaysia

The Asian Conference on Society, Education and Technology 2014
Official Conference Proceedings

Abstract
Forecasting the price of non renewable commodity such as crude oil is a critical task and requires careful attention. Due to the vital role of non renewable commodity in the economics of an organization, forecasting its price has attracted both researchers and practitioners. In this paper, a relatively new Swarm Intelligence (SI) technique, namely Grey Wolf Optimizer (GWO) is utilized as a method for short term time series forecasting. Classified as a nature inspired meta heuristic algorithm, the GWO emerged from the observation of leadership hierarchy and hunting mechanism of grey wolves in the nature. It incorporates of four groups which each of them belong to different level of hierarchy. To date, GWO has been proven to be comparable to existing optimization algorithms, thus carries a great potential for the said time series data. Realized in crude oil price time series data, the efficiency of GWO is measured based on statistical metric viz. Mean Absoluter Percentage Error (MAPE) and is compared against two Evolutionary Computation (EC) algorithms namely Artificial Bee Colony (ABC) and Differential Evolution (DE). Based on the obtained findings, it is noted that the GWO produces the lowest MAPE which is at 5.4779% while the ABC obtained a similar reading at 5.4170%. However, the performance of DE is not at the same par as it produces 11.9320% error rate. Thus, it can be concluded that the GWO may become a competitor in the domain of time series forecasting.

Keywords: Grey Wolf Optimizer, optimization, parameter tuning, time series forecasting,

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 (Zhang & Wei, 2010). 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 (A. Gabralla, Jammazi, & Abraham, 2013; Li & Ma, 2010).

Classified as non renewable natural resources commodity, crude oil is very limited in production and irreplaceable in human time frame (Kemp 2004). 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 (Zhang, Wu, & Zhang, 2010). 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 (Jammazi & Aloui, 2012).

In literature, there are avalanche of studies which present various forecasting techniques for the said time series data. In Jammazi in Aloui (2012), monthly crude oil price forecasting was presented using an improved Back Propagation Neural Network (BPNN). Realized in West Texas Intermediate (WTI) crude oil price, the presented model is compared against conventional BPNN. The findings 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 been demonstrated for crude oil price forecasting (Tehrani & Khodayar, 2011). 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 investigated to forecast crude oil price (A. Gabralla, Jammazi, & Abraham, 2013). In the study, three machine learning algorithms were chosen for comparison purposes which includes 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 (Xu, Zhang, Cheng, Xu, & Zhang, 2014). 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.

Nonetheless, despite the various presented techniques in crude oil price forecasting, finding an effective forecasting models for the said time series data is important. The gaps that exist in the existing works, particularly the Neural Network based model (Jammazi in Aloui, 2012; Tehrani & Khodayar, 2011, Xu, et al., 2014) which is favorably applied in crude oil price forecasting is unavoidable to face with the poor generalization (Cheng, Qian, & Guo, 2006; Xiang & Jiang, 2009) and the requirement of many control parameters to be tuned (Xiang & Jiang, 2009; Zhang, et al., 1998).

In this study, Grey Wolf Optimizer (GWO) (Mirjalili, Mirjalili, & Lewis, 2014) is employed for crude oil price forecasting. As a relatively new Swarm Intelligence (SI) algorithm, GWO is motivated from social behaviour of grey wolves or also known as Canis Lupus which is 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 two important features in GWO which are known as attacking prey and search for prey respectively. This 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 (Mirjalili, et al., 2014). With such performance, the GWO posses a great potential for forecasting the said time series data. This time series data is chosen due to its significant role not only in human life survival but also affect the global economic activities. As to use the GWO in forecasting task, this is done estimating the parameter values, as the ones applied in the existing works (Hadavandi, Ghanbari, & Abbasian-Naghneh, 2010; Mustaffa, Yusof, & Kamaruddin, 2013).

This paper is organized as follows: In Section 2, a brief review on GWO is described. In Section 3, the implemented methodology is presented, followed results and discussion is presented in Section 4. Finally, Section 5 presents the conclusion of the study.

Grey Wolf Optimization

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 on grey wolf population, namely alpha, beta delta and omega. The alpha consists of male and female grey wolf and responsible for decision making such as about hunting, sleeping place and others. Due to its dominant role, they are placed at the top of the hierarchy. As the most dominant pack, it is measures in terms of the best in managing the pack, not the strongest.

The second level namely beta responsible to help alpha in decision making or any other activities of the pack. The beta can be male or female and would be the best candidate for replacement in alpha if one of the alpha passes away or become old. The beta act as an advisor for the alpha in undertaking discipline of the pack.

Meanwhile, the delta, have to submit the solution to alpha and beta but they dominate the omega. This group consist of scouts, sentinels, elders, hunters and caretakers. Lastly, the omega, which ranked last in the hierarchy, plays the role as scapegoat.

B. Mathematical Model and Algorithm

1. Social Hierarchy

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

2. Encircling Prey

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

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

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

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 $\vec{C}$, is calculated as follows:

$$\vec{A} = 2\vec{a}.\vec{r}_1 - \vec{a} \tag{3}$$

$$\vec{C} = 2.\vec{r}_2 \tag{4}$$

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].

3. Hunting

Commonly, the hunting is guided by the alpha. However, both beta and delta might also 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 so far are saved 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:

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

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

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

Initialize the grey wolf population
Initialize a, A and C
Calculate the fitness of each agent
X_α = the best search agent
X_β = the second best search agent
X_δ = the third best search agent
while (t < Max number of iterations)
    for each search agent
        Update the position of the current search agent by equation (3.7)
    end for
    update a, A and C
    Calculate the fitness of all search agents
    Update X_α, X_β and X_δ
    t = t + 1
end while
return X_α

Figure 1: Pseudo Code of GWO

Methodology

This section elaborates steps taken in developing a GWO forecasting model. It starts with data collection and its pre-processing activities. Upon completion of the forecasting model, relevant performance metrics were employed to determine the effectiveness of the proposed model.

A. Research Data and Data Preparation

In this study, real data of West Texas Intermediate (WTI) crude oil prices are considered in the examination. This is due it is the most famous benchmark price (Xiong, Bao & Hu, 2013). The time series data covered is from December 1, 1997 to June 30, 1998 and are freely obtained from Barchart website ("Barchart," 2012). From the sample, 70% is allocated for training purposes while the rest 30% is set for testing.

B. Experiment Setup

The variables assigned to features involved are as tabulated in Table 1. The input arrangement utilized is as suggested in (Malliaris & Malliaris, 2008). The output was the daily spot price of CL one month ahead (21 trading days).

Table 1. Assigning Input and Output Variables

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

C. The Algorithm of GWO for Crude Oil Price Forecasting

In this study, the goal is to minimize the error between the forecast and actual price of the crude oil. For that purpose, the objective function is served by Mean Absolute Percentage Error (MAPE (see section E). The equation of crude oil price forecasting is modified from (Hadavandi, Ghanbari, & Abbasian-Naghneh, 2010) and defined as follows:

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

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.

D. Benchmarking Techniques

In this study, the results from the GWO are compared with the results produced by the following techniques:

i) Artificial Bee Colony Algorithm (ABC)

The ABC algorithm which has been introduced by Dervis Karaboga (Karaboga, 2005) is enlightened from the intelligent foraging behaviour of honey bees swarm. In the algorithm, it is consists of three groups of bees viz. employed bee, onlooker bees and scout bees.

ii) Differential Algorithm (DE)

Introduced by Storn and Price (1997), DE is inspired by the mechanism of natural selection which considered as extension of GA. The difference between DE and GA is, in DE, all possible solutions have an equal chance in evaluation task, while in GA, the chance of updating the solution is depends on fitness value.

E. Performance Evaluation Metric

The performance of the forecasting algorithm is evaluated via statistical evaluation indices namely Mean Absolute Percentage Error (MAPE) (Hyndman & Koehler, 2006) and prediction accuracy (PA). The definitions of these evaluation metrics are shown as follows:

$$MAPE = \frac{1}{N}\left[\sum_{n=1}^{N}\left|\frac{y_n - y(x_n)}{y_n}\right|\right] \tag{17}$$

$$PA = 100\% - (MAPE \times 100) \tag{18}$$

Where $n = 1, 2, \ldots, x$

y_n = actual values

$y(x_n)$ = predicted values/approximate values by predictor models

N =Number of test data

Empirical Results and Discussion

For comparison purposes, the forecasting performance of GWO is compared against the results produced by ABC (Mustaffa, Yusof, & Kamaruddin, 2013) and DE. According to the obtained results in Table 2, the values of α, β γ, δ and ε produced by GWO are 0.8346, 0.1128, 0.1521, 1 and 1 respectively. The combination of this parameters yield 5.4779% of MAPE in testing. Meanwhile, the ABC capable to produced slightly lower MAPE which is 5.4170%. However, based on paired sample T-test, it shows that the statistical level of the difference of the means between GWO and ABC is not significant at 0.05% significance level (see Table 3). By producing high correlation, which is 0.9902, it indicates that the prediction values produced by both techniques move very much in the same patterns. Hence the insignificant difference. On the other hand, the DE is left far behind when the produced MAPE is more than 10%, which is 11.9320%.

Table 2: Comparison of Prediction Techniques for CL 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

Table 3: Significant Test for CL 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 visual performances of results are shown in Figure 1, which plot the actual and forecast value of GWO and the identified competitors from day 103 to day 146 (testing phase). The dashed line represents the actual price, the solid line show the GWO forecast price, while the diamond mark and crossmark represent the forecast value obtained using ABC and DE respectively.

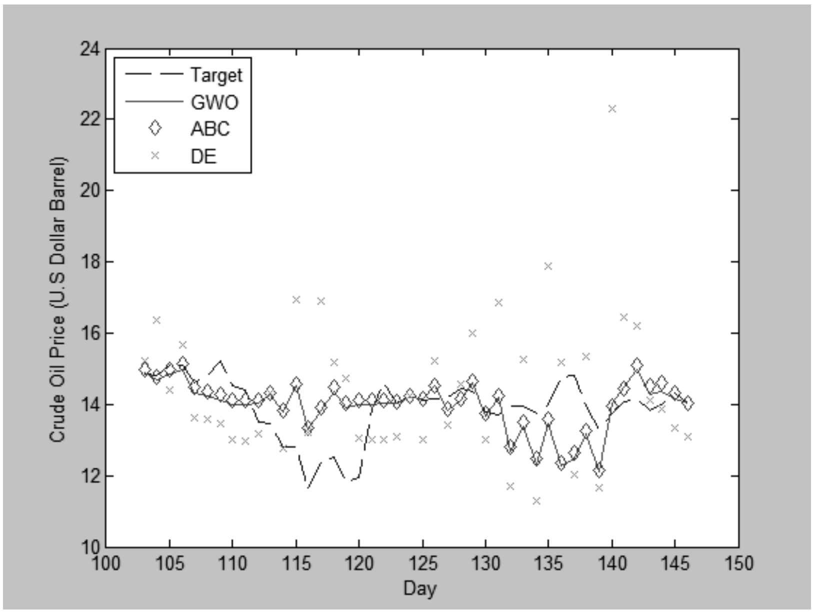

Figure 1: Actual vs. Forecast Values by GWO, ABC and DE

Meanwhile, the actual and forecast values of GWO and the two identified techniques from day 121 to 130 (testing phase) are tabulated in Table 4. The highlighted figures indicated that the respective approach has closer prediction value as compared to the rest in certain days.

Table 4: Actual vs. Forecast Values by GWO, ABC and DE

Day	Actual Price	GWO	ABC	DE
121	13.85	**13.9968**	14.1125	12.989
122	14.59	14.0176	**14.1344**	13.0102
123	14.02	**14.0262**	14.0827	13.1001
124	14.22	14.178	**14.2278**	14.2374
125	14.1	**14.1466**	14.1582	12.9915
126	14.15	**14.5134**	14.537	15.2244
127	14.19	13.8318	**13.8749**	13.422
128	14.45	14.1018	**14.1624**	14.5584
129	14.37	**14.5995**	14.6607	16.0075
130	13.77	13.6482	**13.7291**	13.0227

Conclusion

In this study, a new SI algorithm namely GWO is employed for short term crude oil price forecasting. The efficiency of the proposed technique is measured based on MAPE and PA and is compared against the ones produced by ABC and DE algorithm. Findings of the study reveal competitive results where it is learned that GWO is comparable to ABC algorithm. Thus, it indicates a positive opportunity for future works.

References

A. Gabralla, L., Jammazi, R., & Abraham, A. (2013). *Oil Price Prediction Using Ensemble Machine Learning.* Proceedings of the International Conference on Computing, Electrical and Electronics Engineering (ICCEEE), Khartoum.

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

Cheng, J., Qian, J.-S., & Guo, Y.-N. (2006). Least Squares Support Vector Machines for Gas Concentration Forecasting in Coal Mine. *International Journal of Computer Science and Network Security, 6*(6), 125-129.

Hadavandi, E., Ghanbari, A., & Abbasian-Naghneh, S. (2010, 13-15 Aug. 2010). *Developing a Time Series Model Based on Particle Swarm Optimization for Gold Price Forecasting.* Proceedings of the Third International Conference on Business Intelligence and Financial Engineering (BIFE).

Hyndman, R. J., & Koehler, A. B. (2006). Another Look at Measures of Forecast Accuracy. *International Journal of Forecasting, 22*(2006), 679-688.

Jammazi, R., & Aloui, C. (2012). Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modeling. *Energy Economics, 34*(2012), 828-841.

Karaboga, D. (2005). An Idea Based on Honey Bee Swarm for Numerical Optimization (C. E. D. Engineering Faculty, Trans.): Erciyes University.

Kemp, D. D. (2004). *Exploring Environmental Issues: An Integrated Approach* London: Routledge.

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

Malliaris, M. E., & Malliaris, S. G. (2008). Forecasting Inter-Related Energy Product Prices. *The European Journal of Finance, 14*(6), 453-468.

Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey Wolf Optimizer. *Advance in Engineering Software, 69*(2014), 46-61.

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

Storn, R., & Price, K. (1997). Differential Evolution- a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization, 11*, 341-359.

Tehrani, R., & Khodayar, F. (2011). A hybrid optimized Artificial Intelligent Model to Forecast Crude Oil using Genetic Algorithm. *African Journal of Bussiness Management, 5*(34), 13130-13135.

Xiang, Y., & Jiang, L. (2009, 23-25 Jan. 2009). *Water Quality Prediction Using LS-SVM and Particle Swarm Optimization.* Proceedings of the Second International Workshop on Knowledge Discovery and Data Mining (WKDD)

Xiong, T., Bao, Y., & Hu, Z. (2013). Beyond one-step-ahead forecasting: Evaluation of alternative multi-step-ahead forecasting models for crude oil prices. *Energy Economics, 40*(2013), 405-415.

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

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with Artificial Neural Network: The State of Art. *International Journal of Forecasting, 14*(1998), 35-62.

Zhang, X., Wu, Q., & Zhang, J. (2010). *Crude Oil Price Forecasting Using Fuzzy Time Series.* Proceedings of the 3rd International Symposium on Knowledge Acquisition and Modeling.

Zhang, Y.-J., & Wei, Y.-M. (2010). The Crude Oil Market and the Gold Market: Evidence for Cointegration, Causality and Price Discovery. *Resources Policy, 35*(3), 168-177.