Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science

Procedia Computer Science 91 (2016) 1050 – 1056

Information Technology and Quantitative Management (ITQM2016)

Multi-step-ahead Crude Oil Price Forecasting based on Grey Wave Forecasting Method

Yanhui Chen[a], Yingchao Zou[b,c,*], Yuzhen Zhou[a], Chuan Zhang[d]

[a]School of Economics and Management, Shanghai Maritime University, Shanghai, 201306, China.
[b]College of Information Science and Technology, Beijing University of Chemical Technology, Beijing, 100029, China
[c]School of Economics and Management, Beijing University of Chemical Technology, Beijing, 100029, China
[d]Logistics Engineering College, Shanghai Maritime University, Shanghai, 201306, China

Abstract

Crude oil price plays an important role in the development of global economy. Numerous industry practitioners, researchers and policy makers pay much attention on the fluctuations of crude oil price as while as the accurate prediction of it. This paper uses a graphical prediction method-grey wave forecasting-to forecast multi-step-ahead crude oil price, which enriches the literature of crude oil price prediction and extends the application scope of grey system forecasting theory. The empirical results demonstrate that based on the daily data of crude oil price, grey wave forecasting method performs well in multi-step-ahead prediction and it can also dominate ARMA(1,1).

Peer-review under responsibility of the Organizing Committee of ITQM 2016

Keywords: crude oil price, grey wave forecasting, graphical prediction technology

1. Introduction

As a major raw material of industry and social activities, crude oil price has been widely concerned by numerous industry practitioners, researchers and policy makers. The prediction of crude oil price has already become a hot issue all over the world since unstable crude oil price would introduce turbulence into both crude oil import and export countries, in a further step it would impact global economy negatively. An accurate prediction for crude oil market would help people prepare suitable responses against the disturbance resulted from the fluctuation of crude oil price in advance. Both the rocketing and fast falling of crude oil price block economic development and thence new production development. In particular, a leap in crude oil price would result in an inflation and economy recession in oil-consuming nations, and further negatively impact global

* Corresponding author. Tel.: +86-010-64438793; fax: +86-010-64438793.
E-mail address: evangeline1203@outlook.com.

Peer-review under responsibility of the Organizing Committee of ITQM 2016
doi:10.1016/j.procs.2016.07.147

economy. In contrast, a fast falling of crude oil price would otherwise prohibit the economic development of oil-producing countries, and further generate political instability and social unrest [1, 2]. Therefore, an accurate prediction for crude oil price is pivotal for stable and rapid economic development.

A variety of forecasting models have been formulated for international crude oil price prediction in previous researches. Generally, there are two major strands for the crude oil price forecasting. The first strand can be referred to traditional statistical and econometric techniques, such as linear regression models with explanatory variables, random walk (RW), autoregressive integrated moving average (ARIMA) models, generalized auto regressive conditional heteroskedasticity (GARCH) family models, error correction models (ECM) and regime switching models. For example, Huntington [3] used a sophisticated econometric model with a variety of impact factors to predict crude oil price. Hou and Suardi [4] implemented a nonparametric GARCH model to predict the return volatility in oil price. Lanza, Manera and Giovannini [5] investigated the prices of crude oil and oil products using ECM model. He, Kwok and Wan [6] forecasted WTI prices with daily highs and lows based on a Vector ECM model. Panopoulou and Pantelidis [7] developed two- and three-state regime switching models to forecast crude oil prices. Random walk (RW) and ARIMA are always regarded as the benchmark models in various research articles, for example, Murat and Tokat [8], He, Kwok and Wan [6], Panopoulou and Pantelidis [7] used RW as the benchmark;Yu, Wang and Lai [9], He, Yu and Lai [10], Li, Sun and Wang [11], Yu, Dai and Tang [12] used ARIMA as the benchmark.

However, traditional econometric techniques are always challenged with being incapable of capturing the hidden nonlinear features in crude oil price [12]. In the second strand, artificial intelligence (AI) models with powerful self-learning capacities, such as artificial neural networks (ANNs), support vector machine (SVM) and other intelligent optimization algorithms, have been introduced to forecast crude oil price recently, and the empirical results demonstrated their superiority to traditional methods. For ANN, Abdullah and Zeng [13] introduced ANN to analyze the quantitative data of crude oil price. Azadeh, Moghaddam, Khakzad and Ebrahimipour [14] proposed a flexible algorithm based on artificial neural network (ANN) and fuzzy regression (FR) to forecast long-term oil price. As far as SVM, Xie, Yu, Xu and Wang [15] implemented SVM model to cope with crude oil price forecasting in noisy, uncertain, and complex environments.Khashman and Nwulu [16] also employed SVM to predict crude oil price. Nevertheless the shortcomings of AI models should not be ignored. Time wasting, slow convergence and local minima are listed as the most important disadvantages [12].

A significant point to note is that most of previous researches focused on one-step-ahead forecasting rather than multi-step-ahead forecasting. Xiong, Bao and Hu [17] argued that one-step-ahead prediction provides no information as to the long-term future behavior of crude oil prices and multi-step-ahead forecasting extrapolates the crude oil price series without the availability of outputs in the horizon of interest. They also indicated that multi-step-ahead forecasts of crude oil prices are of greater value to decision-makers in the energy industry than one-step-ahead ones and should be used more widely by practitioners and government agencies in their decision-making related to oil-related investments, risk management and portfolio allocation because they allow for a thorough evaluation of the future behavior of crude oil prices.

This paper uses a kind of graphical prediction technology- grey wave forecasting method- to conduct multi-step-ahead forecasting for daily crude oil price. The primary contribution of this research is the introduction of graphical prediction technology in the analysis of crude oil prices which enriches the literature of crude oil price prediction and extends the application scope of grey system forecasting theory. Grey wave forecasting conducts prediction based on the graph of time series. This method begins with identifying a set of contour lines and establishes GM(1,1) models based on contour time sequences, which are constituted with the intersections of contour lines and time series graph. Grey wave forecasting method conducts prediction without considering the distribution and the stationarity of data[18, 19]. In practice, grey wave forecasting method performs well in forecasting time series with large fluctuation ranges, for example Wan, Wei et al. [20] forecasted Shanghai Composite Index weekly data and found the forecasting wave and the actual wave had the same fluctuating mode. However, compare to the explosive researches on GM(1,1) or GM(1,N)[21-24], few researches applied grey wave forecasting method in time series prediction, not to mention the improvement of grey wave forecasting methods. The representative work involved in this field is Wan, Wei et al. [20]. They chose contour lines according to the crests and troughs of time series graph to fix the limitation that basic (equal-interval) grey wave forecasting method is only effective for time series data with regular fluctuation ranges. However, this method is not efficient since finding all the crests and troughs is a big work load. To overcome the short coming that basic grey wave forecasting method is not suitable for time series with irregular fluctuation range, Chen, Zou and Wang [25] proposed to identify contour lines by using quantiles and filtrate contour time sequences to improve basic (equal-interval) grey wave forecasting method.

The rest of the paper is organized as follows: section 2 provides the brief account of the procedures of grey wave forecasting, base on which section 3 conducts the empirical analysis on crude oil price. Section4 concludes with the summarizing remarks.

2. Methodology

There are three major steps in basic grey wave forecasting: choosing contour lines, identifying contour time sequences, establishing GM(1,1) models based on contour time sequences[18]. The first two steps aim to capture the graphical information of time-series. The last step conducts forecasting with GM(1,1) models, the principle of which is to find the movement rules of the elements in contour time sequences through GM(1,1) modeling. Basic grey wave forecasting method conducts prediction with contour time sequences which include 4 elements or more than 4 elements [18, 20]. But Chen, Zou and Wang [25] suggested when the time series fluctuates severely and the fluctuation ranges are irregular, the numbers of intersections on different contour lines differ greatly. As a result, it is not proper to predict future values based on all the contour time sequences which include 4 elements or more than 4 elements, and it is necessary to filtrate contour time sequences to conduct GM (1, 1) forecasting process. So they proposed to add a new step before the last step and they called it “filtrating contour time sequences”. The new step suggests to divide contour time sequences into two types (qualified and unqualified) according to certain rules. Then the final step establishes GM (1, 1) models based on qualified contour time sequences for in-sample fitting and out-of-sample forecasting, while the unqualified contour time sequences are only used for in-sample fitting. The details of grey forecasting method used in this research will be introduced in the following passages.

In the first step, this paper uses the method proposed by Chen, Zou and Wang [25], that is using quantile to choose unequal-interval contour lines. They suggested the primary reason to choose unequal-interval contour lines is that equal-interval contour lines can't capture the graphical information of time series with irregular fluctuations accurately. This paper takes $X=(x(1),x(2),\cdots,x(n))$ as a time series and takes $X^a=(x^a(1),x^a(2),\cdots,x^a(n))$ as the ascending order of *X*. Then let $\xi_0=x^a(1)$, $\xi_s=x^a(n)$, and let $(\xi_1,\xi_2,\cdots,\xi_{s-1})$ to be the *s*-quantile of the data. This research chooses the *s*+1 horizontal lines decided by $(\xi_0,\xi_1,\xi_2,\cdots,\xi_{s-1},\xi_s)$ as the unequal-interval contour lines.

The second step is identifying contour time sequences. Actually the contour time sequences are constituted by the abscissas of the intersection points between contour lines and graph of times series. Let the original time series to be $X=(x(1),x(2),\cdots,x(n))$, contour line ξ_i intersects with original time series. Let

$X_{\xi_i} = (P_1, P_2, \cdots, P_{m_i})$ to be the set of intersection points, P_j is located on the t_j *th* broken line and the coordinates of P_j are

$$(t_j + \frac{\xi_j - x(t_j)}{x(t_j + 1) - x(t_j)}, \quad \xi_j) . \tag{1}$$

And let

$$q(j) = t_j + \frac{\xi_j - x(t_j)}{x(t_j + 1) - x(t_j)}, \ j = 1, 2, \cdots, m . \tag{2}$$

Thus $Q_i^{(0)} = (q_i(1), q_i(2), \cdots, q_i(m_i))$, $i = 0,1,2,\cdots,s$ is the contour time sequence of contour line ξ_i.

In the third step, this paper introduces the procedure, proposed by Chen, Zou and Wang [25], that is filtrating contour time sequences and dividing contour time sequences into qualified and unqualified contour time sequences. Let t_1^f to be the serial number of first forecasted observation and $Q_i^{(0)} = (q_i(1), q_i(2), \cdots, q_i(m_i))$ .If $t_1^f - q_i(m_i)$ is less than or equal to the threshold *D*, $Q_i^{(0)}$ is the qualified contour time sequence determined by ξ_i. The other contour time sequences are unqualified contour time sequences.

Finally, this paper establishes GM(1,1) models based on contour time sequences. Since this paper filtrates contour time sequences in the previous step, this part just establishes GM(1,1) models based on qualified contour time sequences for both in-sample fitting and out-of-sample forecasting and establishes GM(1,1) models based on unqualified contour time sequences only for in-sample fitting. Specifically, establish GM (1, 1) models based on qualified contour time sequences $Q_i^{(0)} = (q_i(1), q_i(2), \cdots, q_i(m_i))$ to get forecasting values $\hat{q}_i(m_i+1)$, $\hat{q}_i(m_i+2)$, ……, $\hat{q}_i(m_i+k_i)$. Sort all the elements in contour time sequences $Q_0^{(0)}$, $Q_1^{(0)}$, $Q_2^{(0)}$, ……, $Q_s^{(0)}$ by ascending order and delete invalid value. Let the forecasting series is:

$$\hat{q}(1) < \hat{q}(2) < \cdots < \hat{q}(n_s) \tag{3}$$

In which $n_s \le \sum_{i=1}^{s}(m_i + k_i)$. If $\hat{q}(k)$ is on contour line $\xi_{\hat{q}(k)}$, the generated wave through in-sample fitting and out-of-sample forecasting is:

$$X = \hat{X}^{(0)} = \left\{ \xi_{q(k)} + \frac{t - \hat{q}(k)}{\hat{q}(k+1) - \hat{q}(k)} [\xi_{\hat{q}(k+1)} - \xi_{\hat{q}(k)}] \,|\, k = 1, 2, \cdots, n_s \right\} \tag{4}$$

3. Empirical Analysis

3.1. Data

This paper uses daily price data of WTI and Brent crude oil spot prices. The data are obtained from the website of U.S. Energy Information Administration, a central online data warehouse for large number of energy data. Both data sets span from January 4th 2005 to October 30th 2015, the latest date when this research was conducted. Figure 1 indicates that crude oil price fluctuates frequently and the fluctuation ranges are irregular. For example during 2008, crude oil price dropped sharply from more than 140 US dollars per barrel to less than40 US dollars per barrel, and then it climbed to 100 US dollars per barrel in the end of 2010.

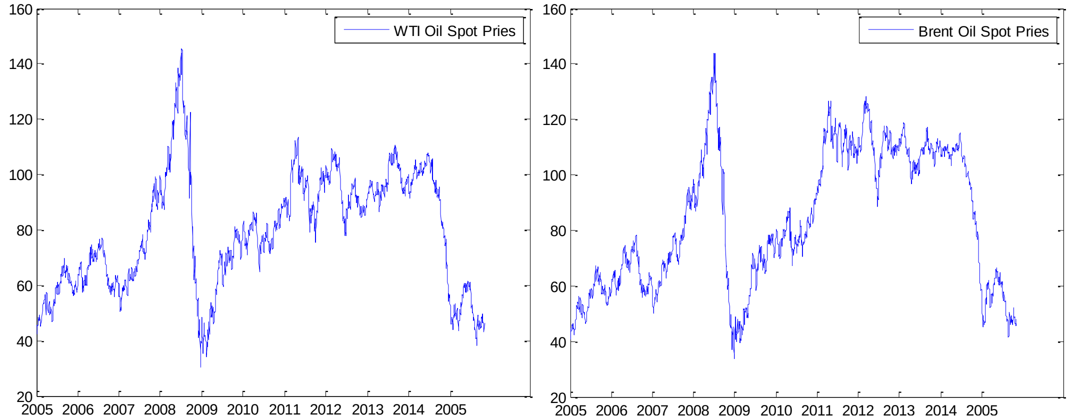

Figure 1 Crude oil spot price from January 4th 2005 to October 30th 2015

The description statistics in Table 1 prove that WTI and Brent oil prices fluctuate severely and fluctuation ranges are irregular, as the maximum is about more than four times of the minimum and the standard deviation is also great. Moreover, according to the result of Jarque-Bera test, we can conclude both data sets don't follow normal distribution which increases the forecasting difficulty when using parametric forecasting models.

Table 1 Description statistics of daily price data of WTI and Brent crude oil

	Mean	max	min	Std. Dev.	Skewness	Kurtosis	Jarque-Bera test	ADF test
WTI	79.261	145.310	30.280	21.004	0.121	2.441	42.152 (0.001)	-0.524 (0.460)
Brent	83.659	143.950	33.730	24.871	0.071	1.707	191.390 (0.001)	-0.424 (0.497)

Table 2 Autocorrelation analysis of WTI and Brent crude oil

	lag	1	2	3	4	5
WTI	AC	0.9956	0.9916	0.9877	0.9836	0.9793
	Prob	0.000	0.000	0.000	0.000	0.000
Brent	AC	0.9971	0.994	0.991	0.988	0.9851
	Prob	0.000	0.000	0.000	0.000	0.000
	lag	6	7	8	9	10
WTI	AC	0.9754	0.9714	0.9675	0.9638	0.9600
	Prob	0.000	0.000	0.000	0.000	0.000
Brent	AC	0.9819	0.9789	0.9759	0.9729	0.9698
	Prob	0.000	0.000	0.000	0.000	0.000

Table 2 presents the autocorrelation coefficients from one lag to ten lags. It indicates that both WTI and Brent crude oil prices are highly auto-correlative, so we can choose contour time sequences, the last elements of which are very close to the serial number of first forecasted observation, as the qualified contour time sequences. So the filtrating rule is based on the autocorrelation characteristic of time series in this paper.

3.2. Forecasting results

This section compares the forecasting results of grey wave forecasting methods with traditional time series forecasting model—ARMA. ARMA model is established based on the logarithm returns of crude oil prices, as the levels of crude oil price are nonstationary. This paper also weighs computation complexity against both Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) which are used to determine the lagged order of ARMA(r,m), as a result an ARMA(1,1) is proved to be the most proper model specification for the data used in this research. This research investigates a 20-step prediction and uses rolling samples to conduct the prediction for 20 times so that we can set a large testing sample (20×20=400) to avoid small sample problems. Particularly, we forecast crude oil prices from October 5^{th} to 30^{th} 2015 (20 trading days) and then forecast the prices from October 2^{nd} to 29^{th} 2015 (also 20 trading days), and on and on.

We use both Root Mean Square Error (RMSE) and Theil's inequality coefficient (TIC) to evaluate forecasting performance. The results are shown in Table 3. They demonstrate that grey wave forecasting method performs better than ARMA(1,1) for both WTI and Brent crude oil price, which suggests grey wave forecasting methods is suitable for multi-step-ahead prediction.

Table 3 Forecasting performance comparison

	Method	RMSE	TIC
WTI	Grey Wave Forecasting	0.2006	0.0052
	ARMA(1,1)	0.2058	0.0054
Brent	Grey Wave Forecasting	0.1987	0.0050
	ARMA(1,1)	0.2228	0.0063

4. Conclusion

This paper applies grey wave forecasting method in the prediction of crude oil price. This research introduces a bran-new mode of forecasting thoughts into crude oil price forecasting. Through the empirical analysis on WTI and Brent crude oil price from January 2005 to October 2015, the results indicate that the grey wave forecasting method can improve the forecasting accuracy and is suitable for multi-step-ahead crude oil price forecasting. In the future, the research can be extended by exploring hybrid grey wave forecasting models to improve forecasting accuracy.

Acknowledgements

This work is supported the National Natural Science Foundation of China (NSFC nos. 71201054, 91224001), the Strategic Research Grant of City University of Hong Kong (No. 7004574), the Fundamental Research

Funds for the Central Universities in BUCT, and the University Yong Teacher Training Program of Shanghai(2015).

References

[1] Gholamian M. R., Fatemi Ghomi S. M. T., Ghazanfari M. A hybrid systematic design for multiobjective market problems: a case study in crude oil markets. Engineering Applications of Artificial Intelligence, 2005, 18(4): 495-509.

[2] Chen Shiu-Sheng, Hsu Kai-Wei. Reverse globalization: Does high oil price volatility discourage international trade? Energy Economics, 2012, 34(5): 1634-1643.

[3] Huntington Hillard G. Oil Price Forecasting in the 1980s: What Went Wrong? The Energy Journal, 1994, 15(2): 1-22.

[4] Hou Aijun, Suardi Sandy. A nonparametric GARCH model of crude oil price return volatility. Energy Economics, 2012, 34(2): 618-626.

[5] Lanza Alessandro, Manera Matteo, Giovannini Massimo. Modeling and forecasting cointegrated relationships among heavy oil and product prices. Energy Economics, 2005, 27(6): 831-848.

[6] He Angela W. W., Kwok Jerry T. K., Wan Alan T. K. An empirical model of daily highs and lows of West Texas Intermediate crude oil prices. Energy Economics, 2010, 32(6): 1499-1506.

[7] Panopoulou Ekaterini, Pantelidis Theologos. Speculative behaviour and oil price predictability. Economic Modelling, 2015, 47(128-136.

[8] Murat Atilim, Tokat Ekin. Forecasting oil price movements with crack spread futures. Energy Economics, 2009, 31(1): 85-90.

[9] Yu Lean, Wang Shouyang, Lai Kin Keung. Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. Energy Economics, 2008, 30(5): 2623-2635.

[10] He Kaijian, Yu Lean, Lai Kin Keung. Crude oil price analysis and forecasting using wavelet decomposed ensemble model. Energy, 2012, 46(1): 564-574.

[11] Li Ziran, Sun Jiajing, Wang Shouyang. An information diffusion-based model of oil futures price. Energy Economics, 2013, 36(518-525.

[12] Yu Lean, Dai Wei, Tang Ling. A novel decomposition ensemble model with extended extreme learning machine for crude oil price forecasting. Engineering Applications of Artificial Intelligence, 2016, 47(110-121.

[13] Abdullah S. N., Zeng X. Machine learning approach for crude oil price prediction with Artificial Neural Networks-Quantitative (ANN-Q) model; proceedings of the 2010 International Joint Conference on Neural Networks (IJCNN), F 18-23 July 2010, 2010.

[14] Azadeh Ali, Moghaddam Mohsen, Khakzad Mehdi, et al. A flexible neural network-fuzzy mathematical programming algorithm for improvement of oil price estimation and forecasting. Computers & Industrial Engineering, 2012, 62(2): 421-430.

[15] Xie Wen, Yu Lean, Xu Shanying, et al. A New Method for Crude Oil Price Forecasting Based on Support Vector Machines. Springer Berlin Heidelberg, 2006.

[16] Khashman A., Nwulu N. I. Intelligent prediction of crude oil price using Support Vector Machines; proceedings of the 2011 IEEE 9th International Symposium on Applied Machine Intelligence and Informatics (SAMI), F, 2011. IEEE.

[17] Xiong Tao, Bao Yukun, Hu Zhongyi. Beyond one-step-ahead forecasting: Evaluation of alternative multi-step-ahead forecasting models for crude oil prices. Energy Economics, 2013, 40(405-415.

[18] Liu Sifeng, Dang Yaoguo, Fang Zhigeng, et al. Grey System Theory and Its Application. 5th ed. Beijing China: Science Press, 2010.

[19] Liu Sifeng, Lin Yi. Grey information: theory and practical applications. Springer Science & Business Media, 2006.

[20] Wan Qin , Wei Yong , Yang Xiongqiong Research on Grey Wave Forecasting Model; proceedings of the 2009 IEEE International Conference on Grey Systems and Intelligent Services(GSIS), F, 2009. IEEE.

[21] Kayacan Erdal, Ulutas Baris, Kaynak Okyay. Grey System Theory-based Models in Time Series Prediction. Expert Systems with Applications, 2010, 37(2): 1784-1789.

[22] Deng Liurui, Hu Zongyi, Yang Shenggang, et al. Improved non-equidistant grey model GM (1, 1) applied to the stock market. Journal of Grey System, 2012, 15(4): 189-194.

[23] He Zhi, Shen Yi, Li Junbao, et al. Regularized multivariable grey model for stable grey coefficients estimation. Expert Systems with Applications, 2015, 42(4): 1806-1815.

[24] Xie Naiming, Liu Sifeng, Yang Yingjie, et al. On novel grey forecasting model based on non-homogeneous index sequence. Applied Mathematical Modelling, 2013, 37(7): 5059-5068.

[25] Chen Yanhui, Zou Yingchao, Wang Shanshan. Grey wave forecasting with qualified unequal-interval contour lines and its application in metal prices prediction; proceedings of the 2015 IEEE International Conference on Grey Systems and Intelligent Services (GSIS), F 18-20 Aug. 2015, 2015. IEEE.