Application of Markov Model in Crude Oil Price Forecasting

Nuhu Isah [1], Abdul Talib Bon [1]

[1] *Universiti Tun Hussein Onn Malaysia*
101 Parit Raja, Batu Pahat, Johor, 86400, Malaysia

Abstract. Crude oil is an important energy commodity to mankind. Several causes have made crude oil prices to be volatile. The fluctuation of crude oil prices has affected many related sectors and stock market indices. Hence, forecasting the crude oil prices is essential to avoid the future prices of the non-renewable natural resources to rise. In this study, daily crude oil prices data was obtained from WTI dated 2 January to 29 May 2015. We used Markov Model (MM) approach in forecasting the crude oil prices. In this study, the analyses were done using EViews and Maple software where the potential of this software in forecasting daily crude oil prices time series data was explored. Based on the study, we concluded that MM model is able to produce accurate forecast based on a description of history patterns in crude oil prices.

Keywords: forecasting; crude oil; price; Markov model.

INTRODUCTION

Crude oil is a complex mixture consisting of more than 200 organic compounds, especially hydrocarbons [1] mostly alkenes and smaller fraction aromatics. Crude oil varies in colour from nearly colourless to tar black, and in viscosity from close to that of water to almost solid. In fact, there are more than 300 different crude types produced around the world [1]. Two of the most important characteristics are density and sulfur content. High-quality cruds are characterized by low density (light) and low sulfur content (sweet) and are typically more expensive than their heavy and sour counterparts: light crudes produce more high-value products, while sweet crude oils require less processing than sour [2].

The oil and gas bearing structure is typically of porous rock such as sandstone or washed out limestone. The sand might have been laid down as desert sand dunes or seafloor. Oil and gas are formed from organic material (tiny plants and animals) deposited in early geological periods (100–200 million years ago) together with sand or silt and later transformed by high temperatures and pressure into hydrocarbons[3]. Formation of oil reservoirs requires meeting of various climatic, geophysical and historical conditions and enormous amount of time.

However, the most oil reach region is Middle East possessing 51 % of total oil reserves, then Central and South America with 16 %. Europe (excluding Eastern Europe and Russia) owns only 1% of world oil reserves [3]. On the country level the biggest oil reserves belong to Saudi Arabia and Venezuela (262.6 and 211.17 bb in 2011 respectively) together owning around one third of world oil. Among ten countries with biggest reserves are also Canada, Gulf states (Iraq, Iran, Kuwait, UAE), Russia, Libya and Nigeria [4, 5, 6].

In the last 21 months, the average price of oil has dropped by about 60 %. Oil prices swing as a result of many factors. Growth in global economic activity can increase demand and drive prices higher, while increasing production rates can lead to decline in prices [7, 8]. Though simple demand and supply theories are useful in describing oil price movements, the factors driving such changes are often difficult to identify. As a result, large fluctuation in oil prices can come as a surprise, as was the case with the recent decline starting in mid-2014 [8].

The current period of price instability is not unique, oil prices also declined by over 50% during global recession of 2007 to 2009 [9]. However, the decline in crude oil price during the recession was due to a noticeable slowdown in global economic activity. Soft global demand also caused prices for goods and services in addition to crude oil to fall suddenly [10]. These movements propose that only a portion of the decline in 2014 is likely due to fragile global economic activity. Increase in supply was another justification for the oil price decline that started in 2014 [11]. Total global oil production increased 3.7% year over year as of December 2014. This increase is on the higher side, though not remarkable by the standards of the past five years.

However, prices are affected not because of the changes in production, but by changes in production relative to what the market projected. Higher production levels in the United States due to new technologies were to some extent expected. But some developments on the supply side were unexpected, such as the comeback of Libyan production, the refusal of the OPEC to reduce supply, and the probable future return of Iran on the global oil market, which likely caused shifts in expectations of future oil supply relative to demand [7].

However, forecasting crude oil price has been one of the biggest challenges to the artificial intelligent (AI) community [8]. The objective of forecasting research has been largely beyond the capability of traditional AI research which has mainly focused on developing intelligent systems that are supposed to emulate human intelligence [7, 8]. By its nature crude oil price is mostly complex, nonlinear and volatile [9]. The rate of price swings in such series depends on many factors such as economic, political and social [9] ets. Therefore, developing AI systems for this kind of forecasting requires an iterative process of knowledge discovery and system improvement through data mining, knowledge engineering, theoretical and data-driven modelling, as well as trial and error experimentation. Crude oil has become an integral part of the global economy [10]. Any fluctuation in crude oil prices affects our personal and corporate financial lives, and the economic health of a country [10]. A bright prediction model for crude oil price forecasting would be highly desirable and would be of wider interest.

A substantial amount of research has been published in recent times and is continuing to find an optimal prediction model for crude oil price [12, 13]. Most of the forecasting research has employed the statistical time series analysis techniques like ARMA model, GARCH model as well as the multiple regression models [14]. In recent years, numerous crude oil price forecasting techniques based on AI, including artificial neural networks (ANN), fuzzy logic, hybridization of ANN and fuzzy system, support vector machines have been proposed [15]. Most of them have their own shortcomings [15]. For example, ANN is very much problem oriented because of its chosen structural design [12]. Some researchers have used fuzzy systems to develop a model to forecast crude oil price behaviour. To build a fuzzy system one requires some background expert knowledge [13]. In this paper, we make use of the well-established Markov Model (MM) to forecast crude oil price. The MM have been extensively used in the area like speech recognition, DNA sequencing, electrical signal prediction and image processing [14] etc. In here, MM is used in a way to develop forecasting model. We locate pattern(s) from the past datasets that match with today's crude oil price behaviour, then interpolate these two datasets with appropriate neighbouring price elements and forecast tomorrow's crude oil price.

METHODOLOGY

If we have a state sequence $\{q_n \cap \in \mathrm{N}^+\}$, we can find the transition frequency F_{ij} in the sequence by counting the number of transitions from state S_i to state S_j in one step. Then the one-step transition frequency matrix for the sequence $\{q_n\}$ can be constructed as follows (1)-(2):

$$F = \begin{bmatrix} F_{11} & F_{12} & F_{13} & \cdots & F_{1m} \\ F_{21} & F_{22} & F_{23} & \cdots & F_{2m} \\ F_{31} & F_{32} & F_{33} & \cdots & F_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ F_{m1} & F_{m2} & F_{m3} & \cdots & F_{mm} \end{bmatrix}, \qquad (1)$$

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2m} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mm} \end{bmatrix}, \quad (2)$$

where $a_{ij} = \dfrac{F_{ij}}{\sum\limits_{j=1}^{m} F_{ij}}, \, if \, \sum\limits_{j=1}^{m} F_{ij} > 0$;

and $\quad a_{ij} = 0, \, if \, \sum\limits_{j=1}^{m} F_{ij} = 0.$

Suppose there are three states: Up, Same and Down, to simulate the movement of the crude oil price. Here is an explanation of Up, Same and Down. We only compare two crude oil prices, v_n is the current price, v_{n-1} is the price of yesterday. If $v_n - v_{n-1} > 0$ is called Up, $v_n - v_{n-1} < 0$ is called Down and $v_n - v_{n-1} = 0$ represents Same.

We use the information to calculate the transition probabilities. There are only three states. On the assumption that the state space is

$S = (S_1, S_2, S_3)$. S_1 = Up, S_2 = Same and S_3 = Down.

Up, Same and Down are three states, which are decided by comparing the previous closing price and the current closing price. We calculate the number of days that both the first day and the second day are up by using the data. We also get the number of days that the first day is down and the second day is up.

RESULTS AND DISCUSSION

We chose the true values of the closing index from the WTI dating from the 2 January, 2015 to 29 May, 2015 yielding 103 trading days. The closing index is chosen to model the process.

In this model, there are three states, on the assumption that the state space is $S = (S_1, S_2, S_3)$. The definition of Up is $u_n - u_{n-1} > 1$, where the u_n is the current closing index and the u_{n-1} is the previous closing index. The definition of Same is $|u_n - u_{n-1}| \leq 1$, the definition of Down is $u_n - u_{n-1} < 1$. We train the true values of the closing index and use the definition of the states to get the Figure 1 and the Table 1.
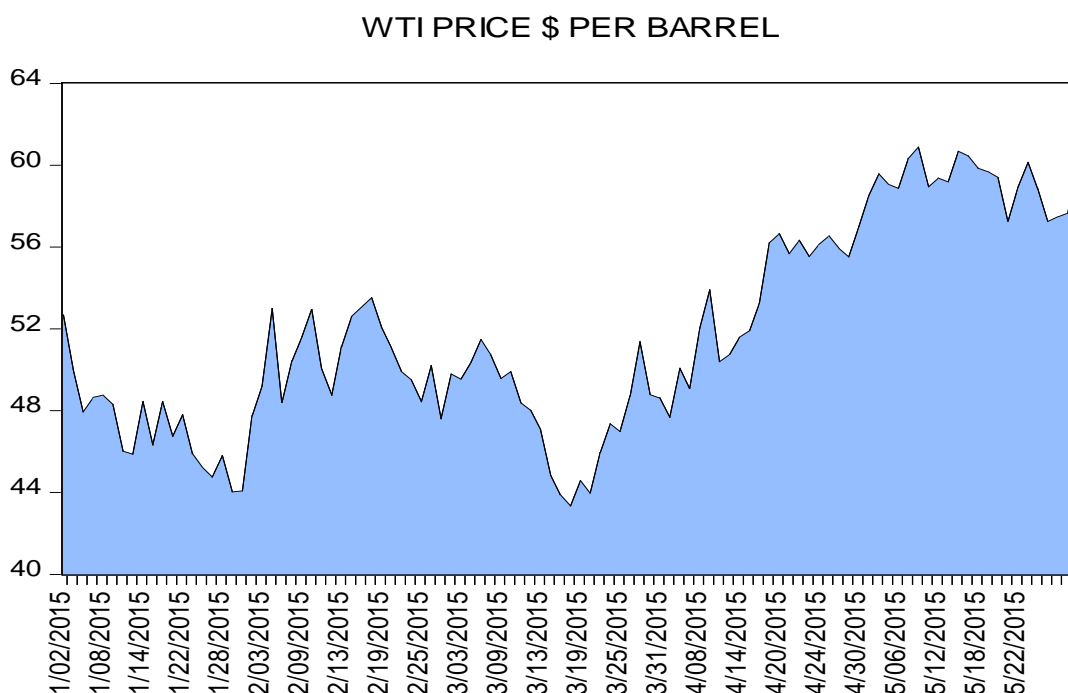
WTI PRICE $ PER BARREL

Figure 1 – Actual movement of the crude oil prices closing index from 2 January to 29 May 2015

Table 1 – States of Transition Matrices

Date	State	Date	State	Date	State	Date	State	Date	State
1/02	Down	2/09	Up	3/16	Down	4/21	Down	5/27	Up
1/05	Down	2/10	Down	3/17	Down	4/22	Up	5/28	Up
1/06	Down	2/11	Down	3/18	Up	4/23	Up	5/29	Up
1/07	Up	2/12	Up	3/19	Down	4/24	Down		
1/08	Up	2/13	Up	3/20	Up	4/27	Down		
1/09	Down	2/16	Down	3/23	Up	4/28	Up		
1/12	Down	2/17	Up	3/24	Down	4/29	Up		
1/13	Down	2/18	Down	3/25	Up	4/30	Up		
1/14	Up	2/19	Down	3/26	Up	5/01	Down		
1/15	Down	2/20	Down	3/27	Down	5/04	Down		
1/16	Up	2/23	Down	3/30	Down	5/05	Up		
1/20	Down	2/24	Down	3/31	Down	5/06	Up		
1/21	Up	2/25	Up	4/01	Up	5/07	Down		
1/22	Down	2/26	Down	4/02	Down	5/08	Up		
1/23	Down	2/27	Up	4/06	Up	5/11	Down		
1/26	Down	3/02	Down	4/07	Up	5/12	Up		
1/27	Up	3/03	Up	4/08	Down	5/13	Down		
1/28	Down	3/04	Up	4/09	Up	5/14	Down		
1/29	Up	3/05	Down	4/10	Up	5/15	Down		
1/30	Up	3/06	Down	4/13	Up	5/18	Down		
2/02	Up	3/09	Up	4/14	Up	5/19	Down		
2/03	Up	3/10	Down	4/15	Up	5/20	Up		
2/04	Down	3/11	Down	4/16	Up	5/21	Up		
2/05	Up	3/12	Down	4/17	Down	5/22	Down		
2/06	Up	3/13	Down	4/20	Up	5/26	Down		

We chose the true values of the closing index of crude oil from WTI dated 2 January to 29 May, 2015 yielding 103 trading days. We define the given transition matrices as Up, Same and Down. In this model, the three states transition matrices are assumed to be $S = (S_1, S_2, S_3)$.

To find the trend of the stock index movement, we need to find the state transition probability by calculating the number of days that both first day and second day are the up, we could find the probability from up to up. Then we to get the number of days that first day is up and second day is down.

$$S_1 \to S_1 \to 23\,days \to S_1 \qquad S_2 \to 0\,days \qquad S_1 \to S_3 \to 26\,days$$

$$S_2 \to S_2 \to 0\,days \to S_2 \qquad S_2 \to 0\,days \qquad S_2 \to S_3 \to 0\,days$$

$$S_3 \to S_1 \to 27\,days \to S_3 \qquad S_2 \to 0\,days \qquad S_3 \to S_3 \to 26\,days$$

Then we get the transition matrix as follows:

$$\hat{A} = \begin{bmatrix} \dfrac{23}{49} & 0 & \dfrac{26}{49} \\ 0 & 0 & 0 \\ \dfrac{27}{53} & 0 & \dfrac{26}{53} \end{bmatrix} \rightarrow \quad \hat{A} = \begin{bmatrix} 0.4694 & 0 & 0.5306 \\ 0 & 0 & 0 \\ 0.5094 & 0 & 0.4906 \end{bmatrix} \rightarrow$$

$$\rightarrow A^2 = \begin{bmatrix} 0.4906 & 0 & 0.5094 \\ 0 & 0 & 0 \\ 0.4890 & 0 & 0.5110 \end{bmatrix} \rightarrow \lim_{n \to \infty} A^n = \begin{bmatrix} 0.4898 & 0 & 0.5102 \\ 0 & 0 & 0 \\ 0.4898 & 0 & 0.5102 \end{bmatrix}$$

Trend of the WTC Crude Oil Price Forecasting from 2 January to 29 May 2015 presents in the Figure 2.
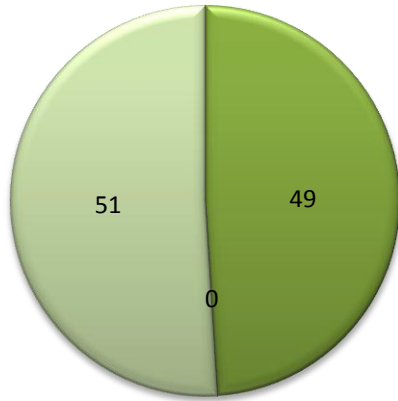

Figure 2 – Trend of the WTC Crude Oil Price Forecasting from 2 January to 29 May 2015, %

In the above model, we got information about three states which are up, same and down. According to the above result, the transition matrix is stable and the most likely trend of index is down, since the probability of down is biggest. The previous price dated May 29 2015 was $60.25 and the price of predicted day dated 1 June 2015 was $60.24 respectively. This shows that the forecasting is tune to be acurate and reliable.

CONCLUSION

This paper explained the application of Markov Model in crude oil price forecasting, crude oil is important natural resources to mankind. The swinging of crude oil prices has affected many economic sectors and stock market indices. In this study, daily crude oil prices data were obtained from WTI dated 2 January to 29 May 2015, the analyses were done using EViews and Maple software where the potential of this software in forecasting daily crude oil prices time series data was explored. Based on the study, we concluded that Markov Model is able to produce accurate forecast based on a description of history patterns in crude oil prices.

ACKNOWLEDGMENT

REFERENCES

1. Xiu, S., & Shahbazi, A. (2012). Bio-oil production and upgrading research: A review. *Renewable and Sustainable Energy Reviews, 16*(7), 4406–4414. doi: 10.1016/j.rser.2012.04.028

2. Xie, W, Yu, L, Xu, S., & Wang, S. (2006). A new method for crude oil price forecasting based on support vector machines. In V. N. Alexandrov, van G. D. Albada, P. M. A. Sloot, J. Dongarra (Eds.), *Computational Science – ICCS 2006. Lecture Notes in Computer Science* (Vol. 3994, p. 444–451). Berlin: Springer Heidelberg. doi: 10.1007/11758549_63

3. Tang, L., & Hammoudeh, S. (2002). An empirical exploration of the world oil price under the target zone model. *Fuel And Energy Abstracts, 24*(6), 577–596. doi: 10.1016/S0140-6701(03)81648-X

4. Radchenko, S. (2005). Oil price volatility and the asymmetric response of gasoline prices to oil price increases and decreases. *Energy economics, 27*(5), 708–730. doi: 10.1016/j.eneco.2005.06.001

5. Pereboichuk, B. (2013). Modeling of Crude Oil Prices With a Special Emphasis on Macroeconomic Factors (Doctoral thesis). Retrieved from http://studenttheses.cbs.dk/bitstream/handle/10417/4420/bogdana_pereboichuk.pdf?sequence

6. Kilian, L., & Murphy, D. P. (2014). The role of inventories and speculative trading in the global market for crude oil. *Journal of Applied Econometrics, 29*(3), 454–478. doi: 10.1002/jae.2322

7. Kaufmann, R. K., Bradford, A., Belanger. L. H., Mclaughlin. J. P., & Miki, Y. (2008). Determinants of OPEC production: Implications for OPEC behavior. *Energy Economics, 30*(2), 333–351. doi: 10.1016/j.eneco.2007.04.003

8. Kaufmann, R. K. (2011). The role of market fundamentals and speculation in recent price changes for crude oil. *Energy Policy, 39*(1), 105–115. doi: 10.1016/j.enpol.2010.09.018

9. Davig, B. T, Nie, J., & Smith, A. L. (2015). Evaluating a Year of Oil Price Volatility. Retrieved from https://www.kansascityfed.org/~/media/files/publicat/econrev/econrevarchive/2015/3q15davigetal.pdf

10. Bopp, A. E., & Lady, G. M. (1991). A comparison of petroleum futures versus spot prices as predictors of prices in the future. *Energy Economics, 13*(4), 274–282. doi: 10.1016/0140-9883(91)90007-m

11. Chatfield, C. (2014). *The analysis of time series: an introduction* (6th ed.). Ontario: Hoboken CRC Press.

12. Teo, T. T., Logenthiran, T., & Woo, W. L. (2016, November). Forecasting of photovoltaic power using extreme learning machine. In *2016 IEEE Region 10 Conference (TENCON)*, Singapore, 2016 (p. 455–458). doi: 10.1109/TENCON.2016.7848040

13. Li, H., Pan, Y., & Zhou, Q. (2015). Filter design for interval type-2 fuzzy systems with D stability constraints under a unified frame. *IEEE Transactions on Fuzzy Systems, 23*(3), 719–725. doi: 10.1109/tfuzz.2014.2315658

14. Farhadi, H., AmirHaeri, M., & Khansari, M. (2015). Alert correlation and prediction using data mining and HMM. *The ISC International Journal of Information Security, 3*(2), 77–101. doi: 10.22042/isecure.2015.3.2.3

15. Wilson, A. D., & Bobick, A. F. (1999). Parametric hidden markov models for gesture recognition. *IEEE transactions on pattern analysis and machine intelligence, 21*(9), 884–900. doi: 10.1109/34.790429