International Journal of Mathematical, Engineering and Management Sciences
Vol. 5, No. 3, 488-496, 2020
https://doi.org/10.33889/IJMEMS.2020.5.3.040

Sugeno Intuitionistic Fuzzy Generator Based Computational Technique for Crude Oil Price Forecasting

Gunjan Goyal
Department of Mathematics,
Jaypee Institute of Information Technology, Noida, India.
E-mail: gunjangoyal.04@gmail.com

Dinesh C. S. Bisht
Department of Mathematics,
Jaypee Institute of Information Technology, Noida, India.
Corresponding author: drbisht.math@gmail.com

(Received November 30, 2019; Accepted February 18, 2020)

Abstract
Crude oil being a significant source of energy, change of crude oil price can affect the global economy. In this paper, a new approach based on the intuitionistic fuzzy set theory has been implemented to predict the crude oil price. This paper presents the intuitionistic fuzzy time series forecasting algorithm to enhance the efficacy of time series forecasting which includes fuzzy c-means clustering to obtain the optimal cluster centers. Further, a computational technique is proposed for the construction of triangular fuzzy sets and these fuzzy sets are converted to intuitionistic fuzzy sets with the help of Sugeno type intuitionistic fuzzy generator. The popular benchmark dataset of West Texas Intermediate crude oil spot price is used for the validation process. The numerical results when compared with existing methods notify that the proposed method enhances the accuracy of the crude oil price forecasts.

Keywords- Intuitionistic fuzzy set, Sugeno type complement function, Fuzzy c-means clustering, Crude oil price forecasting, Fuzzy time series.

1. Introduction

Crude oil price is one of the major factors affecting the world's economy. Forecasting of crude oil prices has been gaining a lot of attention as these forecasts are very useful to industries, investors, government, household expenses and so on. Several prominent researches have been made in this field due to the explosive nature of the oil price. It depends on many factors which makes the price varying from time to time in an unpredictable manner. Crude oil is one of the main input and most traded commodity in the world. The uncertainty in crude oil price affects different sectors in different ways. Crude oil is one of the inputs that cost the firm and variation in its price creates difficulty for the decision maker to take important investment decisions (Henrique and Sadorsky, 2011). Also, it affects the consumers in various situations for instance, when petroleum and gas prices vary or when there is variation in the cost of products whose production requires crude oil. When the price of oil increases, it is in favor of oil producers but against the oil consumers and vice-a-versa (Phan et al., 2015). Thus, the prediction of crude oil prices carries huge importance in the market as well as corporate sectors.

Significant researches are made on the qualitative and quantitative bases in the literature. Xie et al. (2006) proposed a method based on support vector machine (SVM), including pre-processing, training and learning. An adaptive neuro-fuzzy inference system (ANFIS) was also employed for the forecasting process (Rubinstein et al., 2015). A method integrating ensemble empirical mode

International Journal of Mathematical, Engineering and Management Sciences
Vol. 5, No. 3, 488-496, 2020
https://doi.org/10.33889/IJMEMS.2020.5.3.040

decomposition (EEMD), adaptive particle swarm optimization (APSO), and relevance vector machine (RVM) was proposed to predict the price based on decomposition and ensemble framework (Li et al., 2016). Also, using different machine learning techniques a hybrid model was proposed using multi-layer perception, support vector regression (SVR) and Holt Winter exponential smoothing techniques (Kanchymalay et al., 2017) in which it was observed that SVR performs better than other two techniques. A new reconstruction rule based on fuzzy clustering and dynamic time wrapping (DTW) was proposed by Chai et al. (2019) using complete integration empirical mode decomposition algorithm and autoregressive integrated moving average (ARIMA). These traditional techniques lack when there is uncertainty in the historical data due to which Song and Chissom (1993) proposed the concept of fuzzy time series (FTS). After which FTS has been used in various applications (Bisht et al., 2013; Jain and Bisht, 2015; Jain et al., 2017; Jain et al., 2018; Goyal and Bisht, 2019). Zhang et al. (2010) proposed a model to predict the short-term crude oil price using fuzzy time series.

The theory of intuitionistic fuzzy set (IFS) was proposed by Atanassov (1986) as an extension of the fuzzy set (FS) theory to deal with uncertainty and vagueness. In IFS, the non belongingness of an element is also taken in consideration with the belongingness of an element to a set. Kumar and Gangwar (2015) proposed the concept of intuitionistic fuzzy time series combining IFS and FTS. In this paper, an improved method of fuzzy time series forecasting based on IFS theory is proposed for the crude oil price prediction. This method is applied for crude oil price prediction. The methodology starts with the pre-processing of data in which outliers are identified and are replaced by linear interpolation technique. In the process, fuzzy c-means (FCM) clustering is used to obtain the optimal cluster centers which are then used in the process of interval construction with a new computational approach and Sugeno type intuitionistic fuzzy generator (STIFG) is used for the evaluation of non-membership degree. Then hesitancy degree is used for the process of fuzzification and for different order of fuzzy logical relationships (FLRs), the process is evaluated. Also, new rules are formed for the defuzzification process. The model is applied to the benchmark dataset of West Texas Intermediate (WTI) crude oil price and with the help of root-mean squared error (RMSE) and mean average percent error (MAPE), the accuracy of the model is determined.

2. Preliminaries

2.1 Fuzzy c-means Clustering

FCM is a clustering method which allows each data point to belong to more than one cluster with different membership degree which is based on minimization of the objective function J_m, where D is the total number of data points (y_i's), C_j is the center of j^{th} cluster and N is the number of clusters and m is the exponent of fuzzy partition matrix.

$$J_m = \sum_{i=1}^{D}\sum_{j=1}^{N}\mu_{ij}^{m}\left\|y_i - c_j\right\|^2 \tag{1}$$

Following steps are enrolled in performing FCM:

(i) Initially the cluster membership values are randomly generated.
(ii) Cluster centers C_j's are calculated using eq. (2).

$$C_j = \frac{\sum_{i=1}^{D}\mu_{ij}^{m} y_i}{\sum_{i=1}^{D}\mu_{ij}^{m}} \tag{2}$$

International Journal of Mathematical, Engineering and Management Sciences
Vol. 5, No. 3, 488-496, 2020
https://doi.org/10.33889/IJMEMS.2020.5.3.040

(iii) According to eq. (3), these C_j's are updated.

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{N}\left(\frac{y_i - C_j}{y_i - C_k}\right)^{\frac{2}{m-1}}} \tag{3}$$

(iv) Compute J_m using eq. (1).

Steps 2 to 4 are repeated until a termination criterion is satisfied.

2.2 Intuitionistic Fuzzy Set and Intuitionistic Fuzzy Time Series

An intuitionistic fuzzy set (IFS) *I* on universe of discourse Ω is defined as eq. (4) where $\mu_I(y)$ is the membership degree, $\nu_I(y)$ is the non membership degree of data point y.

$$I = \{< y, \mu_I(y), \nu_I(y) >, y \in \Omega, 0 \le \mu_I(y) + \nu_I(y) \le 1\} \tag{4}$$

The value of $\pi_I(y) = 1 - \mu_I(y) - \nu_I(y)$ is known as the degree of non determinacy (hesitancy factor) of y to the IFS *I*. Let IFS $I_s(t)$ be defined on Ω(t) (t=..., 0, 1, 2, ...), then collection of such IFSs is known as Intuitionistic fuzzy time series (IFTS) (Kumar and Gangwar, 2015).

2.3 Sugeno Type Intuitionistic Fuzzy Generator

To construct the IFS from the FS, intuitionistic fuzzy generators are used. Fuzzy generator φ is a type of fuzzy complement satisfying the property given by eq. (5). Sugeno type fuzzy complement is defined as $N(\mu(y))$.

$$\varphi(y) \le 1 - y, \forall\, y \in [0,1] \tag{5}$$

$$N(\mu(y)) = \varphi_\lambda(\mu(y)) = \frac{1-\mu(y)}{1+\lambda\mu(y))}, -1 < \lambda < \infty \tag{6}$$

But for $-1 \langle \lambda \langle 0$, $\varphi_\lambda(\mu(y))$ does not satisfy the property i.e., $\varphi(\mu(y)) \ge 1 - \mu(y)$. Therefore,

$$N(\mu(y)) = \varphi_\lambda(\mu(y)) = \frac{1-\mu(y)}{1+\lambda\mu(y))}, \lambda \ge 0 \tag{7}$$

3. Algorithm of Proposed Method for Forecasting Using IFTS

This section describes the proposed methodology in two phases comprising of pre-processing, fuzzification, construction of FLRs and defuzzification.

Phase I: Pre-processing and Fuzzification

Step 1. Using the generalized extreme studentized deviate (ESD) test (Grubbs, 1950), the data set is processed to find out the outliers in the dataset and is replaced by linear interpolation technique.

Step 2. Now using fuzzy c-means clustering, the dataset is divided into the desired number of clusters after which obtained cluster centers are then used for the construction of intervals. Length of these cluster centers are calculated and further partitioning is done using the following procedure:

(i) Test for outliers is done for a series of the length of cluster centers.
(ii) To remove the outlier length obtained, the average length of the remaining series is calculated.
(iii) Then the length as outlier is divided by average length to obtain q partitioning points and outlier length is equally divided into q parts.

$$q = \frac{Outlier\ length}{Average\ length} \tag{8}$$

(iv) Also, the universal set is defined as Ω = [L, U] where L = lowest value of cluster center – average length and U = highest value of cluster center + average length.
(v) Now from the obtained final series of partitioning points p_k (k=1, 2, ..., n), the intervals v_l (l=1, 2, ..., n-1) are constructed.

Step 3. Assemble the triangular fuzzy sets (TFS) T_s (s=1, 2, ..., n-2) as eq. (9) and membership degree is evaluated using the triangular membership function as eq. (10).

$$T_s = [a, b, c] = [p_s, p_{s+1}, p_{s+2}] \tag{9}$$

where $$\mu_I(y) = \begin{cases} 0 & x \le a \\ \dfrac{x-a}{b-a} & a \le x \le b \\ \dfrac{c-x}{c-b} & b \le x \le c \\ 0 & x \ge c \end{cases} \tag{10}$$

Step 4. Construct IFS I_s corresponding to TFS T_s in which the membership degree is the same as of TFS and non-membership degree is evaluated using Sugeno type intuitionistic fuzzy complement as eq. (11). Therefore, IFS I is defined as $I = \{< y, \mu_I(y), \nu_I(y) >, y \in \Omega\}$.

$$\nu_I(y) = \frac{1-\mu_I(y)}{1+\lambda * \mu_I(y)}, \lambda \ge 0 \tag{11}$$

Step 5. Evaluate the hesitancy factor π of each y_i corresponding to each I_s as eq. (12).

$$\pi_{I_s}(y_i) = \lambda * \mu_{I_s}(y_i) * \nu_{I_s}(y_i) \tag{12}$$

Step 6. On the basis of hesitancy factor π, the data set is fuzzified using following rules:
Rule 1 If y_i belongs to only one IFS (say I_w), then y_i is fuzzified as I_w.

Rule 2 If y_i belongs to more than one IFS, then y_i is fuzzified as I_z for the condition defined in eq. (13).

$$\pi_{I_Z}(y) = \min\{\pi_{I_1}(y), \pi_{I_2}(y), \pi_{I_3}(y), \ldots\} \tag{13}$$

Phase II: Construction of Intuitionistic Fuzzy Logical Relationships and Defuzzification

Step 7. For the construction of intuitionistic fuzzy logical relationships (IFLR), the IFS of time t-2, t-1, t and t+1 are considered i.e., I_{t-2}, I_{t-1}, I_t and I_{t+1}. Hence, the IFLR formed is I_{t-2}, I_{t-1}, $I_t \rightarrow I_{t+1}$. Then on the basis of same antecedent, these IFLR are grouped and rule base is constructed.

Step 8. To evaluate the forecasted output at time t, rules given below are considered assuming intuitionistic fuzzified value at time t-3, t-2, t-1 as I_α, I_β, I_γ:

Rule 1 If I_α, I_β, $I_\gamma \rightarrow I_d$, then forecasted output at time t is T_d where T_d is the TFS corresponding to I_d and the defuzzified value is calculated as

$$cv_d = \frac{a+b+c}{3}.$$

Rule 2 If I_α, I_β, $I_\gamma \rightarrow \omega_d I_d$, $\omega_e I_e$, $\omega_f I_f$, then forecasted output at time t is T_d, T_e, T_f where T_d, T_e, T_f are the TFS corresponding to I_d, I_e, I_f and the defuzzified value is calculated as

$$\frac{\omega_d cv_d + \omega_e cv_e + \omega_f cv_f}{\omega_d + \omega_e + \omega_f}.$$

Rule 3 If I_α, I_β, $I_\gamma \rightarrow \#$, then forecasted output at time t is $T_{\gamma-1}$, T_γ, $T_{\gamma+1}$ where $T_{\gamma-1}$, T_γ, $T_{\gamma+1}$ are the TFS corresponding to $I_{\gamma-1}$, I_γ, $I_{\gamma+1}$ and the defuzzified value is calculated as

$$\frac{cv_{\gamma-1} + cv_\gamma + cv_{\gamma+1}}{3}.$$

4. Numerical Illustration

The proposed method is applied to the popular benchmark dataset for international oil prices, i.e., West Texas Intermediate (WTI). The crude oil monthly price from January 1986 to August 2019 is collected from the site https://www.eia.gov/. The data set is divided into two sets: 70% for the training set (Jan 1986 to Jul 2009) and 30% for the testing test (Aug 2009 to Aug 2019). The dataset is processed under ESD test and outliers found at 1% level of significance are replaced by linear interpolation technique. Next using fuzzy c-means clustering, the 14 cluster centers are obtained as Table 1.

Length between the cluster centers C_{13} and C_{14} is 28.1316 which is found to be an outlier and then partitioning points q = 6 is evaluated using eq. (8) and further partitioning is done. So, the final partition points obtained in universal set Ω = [9.285, 107.625] are 14.300, 17.710, 19.769, 21.774, 25.829, 28.911, 32.607, 36.413, 41.571, 48.954, 58.742, 64.672, 74.479, 79.167, 83.856, 88.545, 93.233, 97.922, 102.610. According to eq. (9), the TFSs so obtained are

International Journal of Mathematical, Engineering and Management Sciences
Vol. 5, No. 3, 488-496, 2020
https://doi.org/10.33889/IJMEMS.2020.5.3.040

Table 1. List of cluster centers obtained using fuzzy c-means clustering

Cluster	Cluster center	Length
C_1	14.300	3.4108
C_2	17.710	2.0583
C_3	19.769	2.0055
C_4	21.774	4.0551
C_5	25.829	3.0815
C_6	28.911	3.6963
C_7	32.607	3.8059
C_8	36.413	5.1578
C_9	41.571	7.3830
C_{10}	48.954	9.7885
C_{11}	58.742	5.9293
C_{12}	64.672	9.8074
C_{13}	74.479	28.1316
C_{14}	102.610	

T_1= (9.285, 14.300, 17.710), T_2= (14.300, 17.710, 19.769)
T_3= (17.710, 19.769, 21.774), T_4= (19.769, 21.774, 25.829)
T_5= (21.774, 25.829, 28.911), T_6= (25.829, 28.911, 32.607)
T_7= (28.911, 32.607, 36.413), T_8= (32.607, 36.413, 41.571)
T_9= (36.413, 41.571, 48.954), T_{10}= (41.571, 48.954, 58.742)
T_{11}= (48.954, 58.742, 64.672), T_{12}= (58.742, 64.672, 74.479)
T_{13}= (64.672, 74.479, 79.167), T_{14}= (74.479, 79.167, 83.856)
T_{15}= (79.167, 83.856, 88.545), T_{16}= (83.856, 88.545, 93.233)
T_{17}= (88.545, 93.233, 97.922), T_{18}= (93.233, 97.922, 102.610)
T_{19}= (97.922, 102.610, 107.625).

Then IFSs are constructed by calculating the membership degree using eq. (10), furthermore while evaluating the non-membership degree using Sugeno type intuitionistic fuzzy complement, different values of λ were used such as 0.5, 1, 10 and 100. Then based on the hesitation degree, the data set is fuzzified and the rule base is constructed. According to the rules explained, the data is forecasted and hence defuzzified output is obtained.

Illustration 1. To predict the price of November 2009, the fuzzified input of August 2009, September 2009 and October 2009 is required i.e., I_{13}, I_{12} and I_{13}. Now, I_{13}, I_{12}, $I_{13} \rightarrow$ #, therefore defuzzified output is $\frac{cv_{12}+cv_{13}+cv_{14}}{3} = 72.634$, where cv_{12} is the centroid of TFS T_{12}, cv_{13} is the centroid of TFS T_{13} and cv_{14} is the centroid of TFS T_{14}.

Illustration 2. To predict the price of February 2018, the fuzzified input of November 2017, December 2017 and January 2018 is required i.e., I_{11}, I_{11} and I_{12}. Now, I_{11}, I_{11}, $I_{12} \rightarrow$ I_{11}, I_{11}, I_{12}, I_{12}, therefore defuzzified output is $\frac{2*cv_{11}+2*cv_{12}}{4} = 61.710$ where cv_{11} is the centroid of TFS T_{11} and cv_{12} is the centroid of TFS T_{12}.

Illustration 3. To predict the price of April 2019, the fuzzified input of January 2009, February 2009 and March 2009 is required i.e., I_{10}, I_{11} and I_{11}. Now, I_{10}, I_{11}, $I_{11} \rightarrow I_{12}$, therefore defuzzified output is $cv_{12} = 65.964$ where cv_{12} is the centroid of TFS T_{12}.

International Journal of Mathematical, Engineering and Management Sciences
Vol. 5, No. 3, 488-496, 2020
https://doi.org/10.33889/IJMEMS.2020.5.3.040

5. Evaluation Parameters

To measure the accuracy of the results of proposed intuitionistic time series forecasting model, RMSE and MAPE are used defined by eq. (14) and (15)

$$RMSE = \sqrt{\frac{\sum_{t=1}^{D}\left(\xi(t) - \zeta(t)\right)^2}{D}} \tag{14}$$

$$MAPE = \sqrt{\frac{\sum_{t=1}^{D}\left|\frac{\xi(t) - \zeta(t)}{\xi(t)}\right|}{D}} \tag{15}$$

where ξ(t) is the actual price of crude oil at time t and ζ(t) is the predicted price at time t and D is the number of data points. Smaller the RMSE and MAPE, higher is the accuracy rate.

6. Experimental Results

The proposed method is modeled to predict the price of the crude oil and the monthly dataset of WTI is used with a total of 404 data points. The proposed model is evaluated on the dataset for different orders of FLRs and different values of λ. In the experiment, it is observed that considering different values of λ does not affect the forecasting process. However, the order of FLRs affects the accuracy of the model. In order to show the superiority of the proposed model, results obtained from the fuzzy time series Chen's model (Chen, 1996) are compared with the results of the intuitionistic fuzzy time series model (proposed model). Table 2 and Table 3 show the comparative results of training and testing dataset and graphical representation is shown in Figure 1.

Table 2. Comparative results of training dataset

MODEL	MSE	RMSE	MAPE
Chen's model	33.71	5.81	0.146
Moving average(order 3)	27.46	5.24	0.088
Exponential smoothing(α=0.7)	35.76	5.97	0.105
Proposed model (order 1)	11.39	3.38	0.07
Proposed model (order 2)	15.96	3.99	0.06
Proposed model (order 3)	7.08	2.66	0.05

Table 3. Comparative results of testing dataset

MODEL	MSE	RMSE	MAPE
Chen's model	57.60	7.59	0.0998
Moving average (order 3)	53.54	7.31	0.0862
Exponential smoothing(α=0.7)	64.72	8.04	0.0928
Proposed model (order 3)	47.127	6.864	0.0845

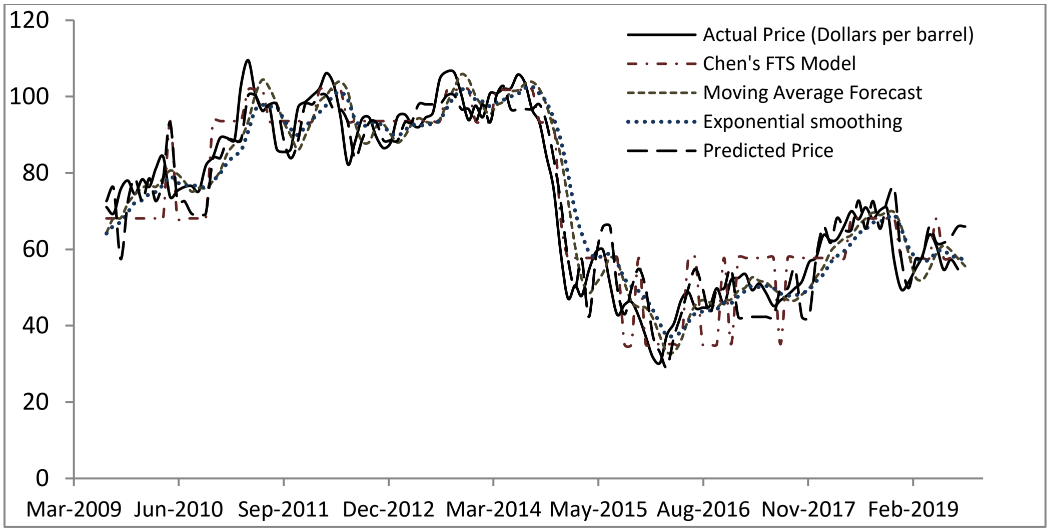

Figure 1. Graphical representation of testing dataset

7. Conclusion

In this paper, a new approach based on intuitionistic fuzzy time series has been proposed to predict the price of crude oil. The proposed model is explained in two phases, firstly pre-processing and fuzzification in which FCM clustering is used with a new computational approach to construct the TFS and STIFG is used for the evaluation of non-membership degree. Then hesitancy degree is used for the process of fuzzification. Second is the construction of rule base and defuzzification in which the proposed model is carried out for first order FLR, second order FLR and third order FLR separately and new rule are formed for the defuzzification process. The model is applied on the benchmark dataset of WTI crude oil price and with the help of evaluation parameters RMSE and MAPE; it is observed that proposed intuitionistic fuzzy set theory based model forecasted better results in comparison to the model based on the fuzzy set theory and moving average method. Further, this work can be extended using the different intuitionistic fuzzy generator for the conversion of fuzzy set to intuitionistic fuzzy set and nature-inspired optimization can be employed to enhance the process.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this work.

Acknowledgments

The authors are grateful to the editor and reviewers for their helpful suggestions.

Reference

Atanassov, K.T. (1986). Intuitionistic fuzzy sets. *Fuzzy Sets and Systems*, *20*(1), 87–96.

Bisht, D., Jain, S., & Raju, M.M. (2013). Prediction of water table elevation fluctuation through fuzzy logic & artificial neural networks. *International Journal of Advanced Science and Technology*, *51*, 107–120.

International Journal of Mathematical, Engineering and Management Sciences
Vol. 5, No. 3, 488-496, 2020
https://doi.org/10.33889/IJMEMS.2020.5.3.040

Chai, J., Wang, Y., Wang, S., & Wang, Y. (2019). A decomposition–integration model with dynamic fuzzy reconstruction for crude oil price prediction and the implications for sustainable development. *Journal of Cleaner Production*, *229*, 775–786.

Chen, S.-M. (1996). Forecasting enrollments based on fuzzy time series. *Fuzzy Sets and Systems*, *81*(3), 311–319.

Goyal, G., & Bisht, D.C. (2019, January). An aggregated higher order fuzzy logical relationships technique. In *AIP Conference Proceedings*, (Vol. 2061 No 1, pp. 020023). AIP Publishing.

Grubbs, F.E. (1950). Sample criteria for testing outlying observations. *The Annals of Mathematical Statistics*, *21*(1), 27–58.

Henriques, I., & Sadorsky, P. (2011). The effect of oil price volatility on strategic investment. *Energy Economics*, *33*(1), 79–87.

Jain, S., & Bisht, D. (2015). Genetic algorithms based fuzzy time series prediction for water table elevation fluctuation. *Aloy Journal of Soft Computing and Applications*, *3*(1), 14-23.

Jain, S., Bisht, D.C., Singh, P., & Mathpal, P.C. (2017, October). Real coded genetic algorithm for fuzzy time series prediction. In *AIP Conference Proceedings*, (Vol. 1897 No. 1, pp. 020021) AIP Publishing.

Jain, S., Mathpal, P.C., Bisht, D., & Singh, P. (2018). A unique computational method for constructing intervals in fuzzy time series forecasting. *Cybernetics and Information Technologies*, *18*(1), 3–10.

Kanchymalay, K., Salim, N., Sukprasert, A., Krishnan, R., & Hashim, U.R. (2017). Multivariate time series forecasting of crude palm oil price using machine learning techniques. In *IOP Conference Series: Materials Science and Engineering*. *2017 International Research and Innovation Summit*, (Vol. 226, No. 1, p. 012117). IOP Publishing. Melaka, Malaysia.

Kumar, S., & Gangwar, S.S. (2015). Intuitionistic fuzzy time series: An approach for handling nondeterminism in time series forecasting. *IEEE Transactions on Fuzzy Systems*, *24*(6), 1270–1281.

Li, T., Zhou, M., Guo, C., Luo, M., Wu, J., Pan, F., Tao, Q., He, T. (2016). Forecasting crude oil price using EEMD and RVM with adaptive PSO-based kernels. *Energies*, *9*(12), 1014.

Phan, D.H.B., Sharma, S.S., & Narayan, P.K. (2015). Oil price and stock returns of consumers and producers of crude oil. *Journal of International Financial Markets, Institutions and Money*, *34*, 245–262.

Rubinstein, S., Goor, A., & Rotshtein, A. (2015). Time series forecasting of crude oil consumption using neuro-fuzzy inference. *Journal of Industrial and Intelligent Information*, *3*(2), 84–90.

Song, Q., & Chissom, B.S. (1993). Fuzzy time series and its models. *Fuzzy Sets and Systems*, *54*(3), 269–277.

Xie, W., Yu, L., Xu, S., & Wang, S. (2006, May). A new method for crude oil price forecasting based on support vector machines. In *International conference on computational science,* (pp. 444-451). Springer, Berlin, Heidelberg.

Zhang, X., Wu, Q., & Zhang, J. (2010). Crude oil price forecasting using fuzzy time series. *2010 Third International Symposium on Knowledge Acquisition and Modeling,* (pp. 213–216). IEEE.