Menemui Matematik (Discovering Mathematics)
Vol. 42, No. 2: 96-104 (2020)

Interval Based in Fuzzy Sliding Window for Forecasting Crude Palm Oil Prices

Nur Fazliana Rahim[1], Mahmod Othman[2] and Rajalingam Sokkalingam[3]
[1]*Centre for Pre University Studies, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak*
[2]*Fundamental and Applied Sciences Department, Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Perak*
[1]rnfazliana@unimas.my, [2]mahmod.othman@utp.edu.my & [3]raja.sokkalingam@utp.edu.my

ABSTRACT

Interval is the main component in time series forecasting, hence a Fuzzy Sliding Window Forecasting Method (SWM) suggested in obtaining intervals of forecasting in the Fuzzy Time Series (FTS). Formerly, almost all the intervals were calculated using class frequency. The intervals are then regrouping into the sub-intervals using the provided category. Whereas in this study, the prediction of interval obtained by embedding the idea of SWM into FTS forecasting. The intention of this suggested method is to further improve the success of a time series forecast and indirectly increase forecasting precision. The daily prices of Crude Palm Oil (CPO) data are taken for verification purposes. Hence, the precision of the suggested method is differentiating the existing forecasting method. The outcome of this method is compared to the other methods and it reveals that the suggested method produces precise intervals determination. The discovery of this study can be used as a replacement of existing forecasting method to get an improved prediction interval.

Keywords: Forecasting, Fuzzy Time Series, Intervals, Sliding Window

INTRODUCTION

Forecasting is an expected value whereby used past and present data. Several methods are used to produce estimation values using a time series study (Ahmad, Ping and Mahamed, 2014). Forecasting has been accomplished using various kinds of techniques (Karia, Bujang and Ahmad, 2013), (Dani and Sharma, 2013) that is applied to solve the forecasting issue, where some of the concepts are adjusted from (Song and Chissom, 1993) and referred to the theory in (Khaliq and Ahmad, 2010). Three steps involve in Fuzzy theory that are fuzzification, determination of rules and fuzzy inference, and defuzzification (Asklany, et. al, 2011).

One of the interesting topics to look into time series is the Fuzzy Sliding Window Forecasting Method (SWM). SWM is introduced by (Datar, et. al, 2002), and it is used in the analysis of time series and ideal in different areas of forecasting (Ben Yahmed, et. al, 2015), (D'Arcy, et. al, 2002), (Kapoor and Bedi, 2013), (Arasu and Manku, 2004). It is evident that the SWM is appropriate in forecasting weather circumstances and the model provides an accurate and precise average forecasting (Kapoor and Bedi, 2013). The procedure of the SWM dictates a point to separate the intervals (Bingham, et. al, 2006). The study in (Rao, et. al, 2015) uses the SWM targeted on climate change and the outcome indicates higher percent accuracy of the method. The study in (Vamitha, et. al, 2012) indicated that mixed FTS with different models able to accomplish good results in forecasting. It is mention in (Dani and Sharma, 2013), SWM is an effective segmented time series forecasting model that should be brought into thought.

In the previous study, all the researchers used the limited intervals for the FTS forecasting approaches and based on expertise suggestion. The interval lengths are importance in forecasting performance and it was stated in (Huarng, 2011), where the study used mean and distribution to find the intervals. In (Huarng and Yu, 2006), researchers are recommended to use the ratios, rather than the same lengths of the interval which able to correctly identify the interval.

However, there are also a study emphasizes that there is still a weakness and drawback to solving for more efficiency and better model to forecast (Kapoor and Bedi, 2013). Therefore, a particular method should be used to provide more precise measurements so that it can be easy to determine the intervals. Research in (Nor,et. al, 2017) used Sliding Window method for rainfall forecasting. The concept of the method was applied to this research in order to determine the class intervals.

Due to the weakness of the general factor, this study introduced a method that can systematically determine the intervals in the forecasting of FTS to reduce forecast errors. The interval obtained is used for forecasting the price of Crude Palm Oil (CPO) to verifying the suggested method. This research was compared to existing forecasting method in (Othman and Azahari, 2016) since the interval used by the method was based on ratio and frequency, while this research used Sliding Window Method to obtain the class intervals.

METHODOLOGY

The procedure used to reach the goals is highlighted in this section. Below is the detailed clarifications that have been done in this study.

Selection of the Data

Daily Crude Palm Oil (CPO) prices with referred on daily historical data of CPO prices are selected for the last half-decade starting the year 2012 to 2016. The highest and lowest prices data are defined for each year.

Fuzzy Sliding Window Forecasting Method

The study starts with determining the universe, U by using daily CPO price data. Table 1 depicts the universe, U for the year 2012 to 2016 based on the data selection.

Table 1: Universe, U

Year	High (RM)	Low (RM)	Universe, U
2012	628.70	395.20	[395.20, 628.70]
2013	580.90	428.40	[428.20, 580.90]
2014	612.10	415.40	[415.40, 612.10]
2015	472.70	352.20	[352.20, 472.70]
2016	596.90	408.60	[408.60, 596.90]
AVERAGE	628.70	352.20	[352.20, 628.70]

The universe, U is partitioned to the same interval size which acquires from the SWM. In a prior study, the 2 week period next to the last year is used (Kapoor and Bedi, 2013) while in this study, a 12x1 matrix present year and 24x1 matrix previous year is used to create the sliding window. Present year matrices consist of data for the year 2016, while the matrix of the previous year consists of data for the year 2014 to 2015.

13 sliding windows are constructed, which similar to the present year matrix size using the size of the matrix in the year before. For each sliding window, the Euclidean distance, Ed_i and mean of Ed_i is computed using Equation (1) and (2), respectively.

$$Ed_{i.j} = \sqrt{\left(x_i - x_j\right)^2} \tag{1}$$

$$\text{Mean of } Ed_i = \frac{\sum Ed_i}{n} \tag{2}$$

where Ed_i refer to Euclidean distance, x_i denotes the data for the present year matrix and x_j denotes the data for the previous year matrix. Below is an example to calculate the Euclidean Distance, Ed_i, and Mean of Ed_i.

$$\begin{aligned} Ed_{i.j} &= \sqrt{(x_i - x_j)^2} \\ Ed_1 &= \sqrt{(421.40 - 535.10_j)^2} \\ &= 113.69 \end{aligned}$$

$$\begin{aligned} \text{Mean of } Ed_i &= \frac{\sum Ed_i}{n} \\ \text{Mean of } Ed_1 &= \frac{\sum 785.76}{12} \\ &= 65.48 \end{aligned}$$

Next, Euclidean distance, Ed_i of which the data with the lowest mean value is selected from the sliding window, which is in W1. Then, referring to this value, the variation and the mean value of variation are computed. The mean-variance labelled as MV1. At the same time, the present year's variance is calculated and labelled MV2. The variations are determined using Equation (3) below.

$$V = AR_t - AR_{t-1} \tag{3}$$

where V represents the variation, AR_t is the data in the present month, while AR_{t-1} is the data in the month before. By using Equation (4), the mean of variation can be found as expressed below (n refer to the number of months).

$$Mean\ of\ Var. = \left| \frac{\sum \text{Variation of each month}}{n} \right| \tag{4}$$

The following calculation shows how to find MV1 and MV2.

$$\begin{aligned} V_p &= AR_t - AR_{t-1} \\ &= 560.72 - 535.10 \\ &= 25.63 \end{aligned}$$

$$\begin{aligned} Mean\ of\ Var_{previous} &= \left| \frac{\sum \text{Variation of each month}}{n} \right| \\ &= \left| \frac{25.63 + 31.51 \_ ... + 22.45}{12} \right| \\ &= 6.98 \end{aligned}$$

$$\begin{aligned} V_c &= AR_t - AR_{t-1} \\ &= 475.80 - 463.15 \\ &= 12.65 \end{aligned}$$

$$\begin{aligned} Mean\ of\ Var_{current} &= \left| \frac{\sum \text{Variation of each month}}{n} \right| \\ &= \left| \frac{12.65 + 34.84 \_ ... + (-46.39)}{12} \right| \\ &= 5.18 \end{aligned}$$

Then, the interval, IV is calculated using the following Equation (5).

$$IV = \left| \frac{\text{Mean Var. Current} + \text{Mean Var. Previous}}{2} \right| \tag{5}$$

$$\begin{aligned} &= \left| \frac{5.18 + 6.98}{2} \right| \\ &= 6.08 \\ &\approx 6 \end{aligned}$$

Generate Forecasts Values

The forecasting rule is classified as in Table 2 referred to (Rahim, et. al, 2018). In this part, the rules used to dictate the CPO prices forecast trend depending on the regulation of the terms and its conditions.

Table 2: Forecasting Rules

Type	Term and Condition
Rule 1	If the highest value of min[max(μ_{A1}×A1, μ_{A2}×0.21A2, μ_{A3}×0.04A3), max(μ_{B1}×B1, μ_{B2}×0.2B2, μ_{B3}×0.27B3), max(μ_{C1}×C1, μ_{C2}×0.25C2, μ_{C3}×0.64C3)], then the CPO prices forecasting go upward at the 0.25 point of the corresponding sub-interval.
Rule 2	If the highest value of min[max(μ_{A1}×0.27A1, μ_{A2}×A2, μ_{A3}×0.78A3), max(μ_{B1}×0.63B1, μ_{B2}×B2, μ_{B3}×0.99B3), max(μ_{C1}×C1, μ_{C2}×C2, μ_{C3}×0.94C3)], then the CPO prices forecasting be the middle of the corresponding sub-interval.
Rule 3	If the highest value of min[max(μ_{A1}×0A1, μ_{A2}×0.12A2, μ_{A3}×A3), max(μ_{B1}×0B1, μ_{B2}×0.27B2, μ_{B3}×B3), max(μ_{C1}×0C1, μ_{C2}×0.34C2, μ_{C3}×C3)], then the CPO prices forecasting go downward at the 0.75 points of the corresponding sub-interval.

When these data are categorized according to the forecasting rule and trend, the distribution of CPO prices is beginning to forecasts.

Performance Measure

The comparative result of the suggested method and the existing FTS method was done in this phase. All forecasts error for CPO prices are tested with Mean Square Error (MSE) and Root Mean Square Error (RMSE). These measures are able to prove the precision of the suggested method and the value can be used to compare the actual data and forecast values (Jolliffe and Stephenson, 2012). It is interpreted as follows.

$$MSE = \left| \frac{\frac{1}{n}\sum_{i}^{n}\left(\text{Actual Data}_{\text{i}} - \text{Forecasted Data}_{i}\right)^{2}}{\text{Actual Data}_{i}} \right| \tag{6}$$

$$RMSE = \sqrt{\frac{\sum_{i}^{n}\left(\text{Actual Data}_{\text{i}} - \text{Forecasted Data}_{i}\right)^{2}}{n}} \tag{6}$$

DISCUSSION AND ANALYSIS

The outcome of the numerical analysis run in this study is revealed and further, discuss in detailed the results. Table 3 shows sliding window samples that calculate Ed_i and Mean of Ed_i. Table 4 depicted the variation and mean variation of the sliding window with the lowest mean value of Euclidean distance, *Edi* while Table 5 shows the calculated variation and mean variation for the present year. From this calculation, there are 6 intervals obtained. Each Interval is split into sub-interval, which using the same steps in the SWM. Table 6 below listed the intervals and sub-intervals.

Table 3: Sliding Window

W1		W2			W13	
Price	Ed_1	Price	Ed_2		Price	Ed_{13}
535.10	113.69	560.73	139.32		451.33	29.92
560.73	84.92	592.24	116.44		456.11	19.69
592.24	85.52	565.61	58.89		445.89	60.83
565.61	4.71	541.45	19.44		437.60	123.29
541.45	0.97	521.66	20.77		448.87	93.47
521.66	3.27	531.41	13.02	…………	455.39	63.00
531.41	22.25	495.22	13.95		432.35	76.81
195.22	53.35	452.56	95.91		387.01	161.55
452.56	119.22	450.77	121.10		395.53	176.35
450.77	79.33	449.15	80.95		417.23	112.88
449.15	81.81	428.88	102.08		390.47	140.50
428.88	136.62	451.33	114.17		403.56	161.94
Mean	65.48	Mean	74.67		Mean	101.69

Table 4. Variation and Mean Variation (Lowest Mean of Edi)

W1		VP
Price	Ed_1	Variation
535.10	113.69	25.63
560.73	84.92	31.51
592.24	85.52	-26.63
565.61	4.71	-24.16
541.45	0.97	-19.79

521.66	3.27	9.76
531.41	22.25	-36.19
495.22	53.35	-42.66
452.56	119.32	-1.79
450.77	79.33	-1.62
449.15	81.81	-20.27
428.88	136.62	22.45
Mean	65.48	6.98

Table 5. Variation and Mean Variation (Present Year)

Current Year, CY		VC
Month	Price	Variation
Jan-16	463.15	12.65
Feb-16	475.80	34.84
Mar-16	510.64	20.10
Apr-16	530.74	-6.07
Mei-16	524.67	3.02
Jun-16	527.69	-4.82
Jul-16	522.87	1.46
Aug-16	524.33	21.03
Sep-16	545.36	5.31
Oct-16	550.67	5.22
Nov-16	555.89	15.79
Dec-16	571.68	-46.39
Mean	525.29	5.18

Table 6: Sliding Window

Interval, *Ui*	**Crude Palm Oil Prices**					**New Sub-interval, Sj**
	2012	**2013**	**2014**	**2015**	**2016**	
U1,1	[395 , 404.75]	[428 , 437.75]	[415 , 424.75]	[352 , 361.75]	[409 , 418.75]	S1
U1,2	[404.75 , 414.5]	[437.75 , 447.5]	[424.75 , 434.5]	[361.75 , 371.5]	[418.75 , 428.5]	S2
U1,3	[414.5 , 424.25]	[447.5 , 457.25]	[434.5 , 444.25]	[371.5 , 381.25]	[428.5 , 438.25]	S3
U1,4	[424.25 , 434]	457.25 , 467]	[444.25 , 454]	[381.25 , 391]	[438.25 , 448]	S4
U2,1	[434 , 441.8]	[467 , 474.8]	[454 , 461.8]	[391 , 398.8]	[448 , 455.8]	S5
U2,2	[441.8 , 449.6]	[474.8 , 482.6]	[461.8 , 469.6]	[398.8 , 406.6]	[455.8 , 463.6]	S6
U2,3	[449.6 , 457.4]	[482.6 , 490.4]	[469.6 , 477.4]	[406.6 , 414.4]	[463.6 , 471.4]	S7
U2,4	[457.4 , 465.2]	[490.4 , 498.2]	[477.4 , 485.2]	[414.4 , 422.2]	[471.4 , 479.2]	S8
U2,5	[465.2 , 473]	[498.2 , 506]	[485.2 , 493]	[422.2 , 430]	[479.2 , 487]	S9
U3,1	[473 , 479.5]	[506 , 512.5]	[493 , 499.5]	[430 , 436.5]	[487 , 493.5]	S10
U3,2	[479.5 , 486]	[512.5 , 519]	[499.5 , 506]	[436.5 , 443]	[493.5 , 500]	S11
U3,3	[486 , 492.5]	[519 , 525.5]	[506 , 512.5]	[443 , 449.5]	[500 , 506.5]	S12
U3,4	[492.5 , 499]	[525.5 , 532]	[512.5 , 519]	[449.5 , 456]	[506.5 , 513]	S13
U3,5	[499 , 505.5]	[532 , 538.5]	[519 , 525.5]	[456 , 462.5]	[513 , 519.5]	S14
U3,6	[505.5 , 512]	[538.5 , 545]	[525.5 , 532]	[462.5 , 469]	[519.5, 526]	S15
U4,1	[512 , 531.5]	[545 , 564.5]	[532 , 551.5]		[526 , 545.5]	S16
U4,2	[531.5 , 551]	[564.5 , 584]	[551.5 , 571]		[545.5 , 565]	S17
U5,1	[551 , 564]		[571 , 584]		[565 , 578]	S18
U5,2	[564 , 577]		[584 , 597]		[578 , 591]	S19
U5,3	[577 , 590]		[597 , 610]		[591 , 604]	S20
U6,1	[590 , 603]					S21
U6,2	[603 , 616]					S22
U6,3	[616 , 629]					S23

Table 7 is depicted some of the forecasted CPO price outcomes. Table 7 shows that none forecasting value of the data for the first month (January). This is due to the previous year's data from is being unconsidered in this study. Based on Table 4, the values of forecasting obtained is closer to that actual price. Next, some forecast values using the existing FTS method as in (Othman and Azahari, 2016) are depicted such Table 8.

Table 7. The Crude Palm Oil (CPO) Prices Forecasting using the suggested method

Actual Price (RM)	**Rule**	**Trend**	**Forecast Value (RM)**
531.21			
527.50	Rule 1	Middle	506.75
568.43	Rule 1	Middle	565.50
618.32	Rule 2	Middle	637.50
591.28	Rule 2	Downward	608.25
553.24	Rule 3	Middle	572.50
568.60	Rule 3	Middle	580.50
558.52	Rule 2	Middle	542.50
513.61	Rule 3	Middle	536.75
421.28	Rule 3	Middle	404.38
431.05	Rule 3	Middle	414.13
411.35	Rule 2	Middle	404.63

Table 8. Forecast values of CPO using existing Fuzzy Time Series method

Actual Price (RM)	Rule	Trend	Forecast Value (RM)
531.21			
527.50	Rule 1	Middle	571.63
568.43	Rule 1	Middle	613.55
618.32	Rule 2	Middle	664.25
591.28	Rule 2	Middle	654.75
553.24	Rule 3	Middle	599.90
568.60	Rule 3	Middle	605.50
558.52	Rule 2	Downward	519.90
513.61	Rule 3	Middle	551.88
421.28	Rule 3	Middle	472.50
431.05	Rule 3	Middle	392.50
411.35	Rule 2	Middle	459.50

As seen in Table 8, the forecast values are quite different compared to the actual values. This means that the existing forecasting method unable to reduce so much forecasting error. Table 9 below compiles the value of errors and the percent of precision obtained using both methods.

Table 9. Performance Measure

Method	MSE	RMSE	Percent Accuracy
Suggested method	0.62	17.51	96.48%
Existing FTS method (Rahim, et. al, 2018)	4.51	46.93	90.42%

Table 9 shows the MSE and RMSE values for the suggested method, which are 0.62 and 17.51. Refer to that value of percent accuracy; it indicates that the suggested method is more accurate than the existing FTS method in (Othman and Azahari, 2016). Therefore, the suggested method can be used as a replacement of the existing forecasting method to obtain intervals of the universe for a better time series forecasting. Figure 1 illustrates the differences between the methods used to forecasts for the year 2012 to 2016.

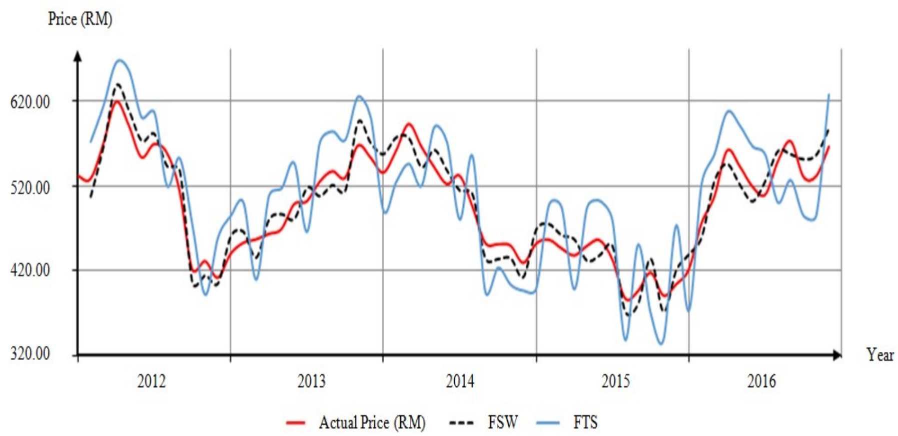

Figure 1: Difference between the actual CPO price and both forecasting method

Referring to Figure 1, the actual CPO price is labeled with the red line. The dashed line (FSW) refers to the suggested method and the existing forecasting method in (Othman and Azahari, 2016) is labeled with the blue line (FTS). Based on the graph, it is pointed out that the suggested method has a significant accuracy of CPO prices forecasting values against the actual CPO price values. This means that the suggested method has the competency to obtain better intervals to accurately execute the CPO prices forecast.

CONCLUSION

Sliding Window Forecasting Method (SWM) is suggested and implemented in forecasting time series data of CPO prices. First, the actual price of CPO was selected and defining the universe for each year 2012 to 2016. By executing the SWM, the intervals are obtained and the CPO prices forecasting is compared between the suggested method and the existing Fuzzy Time Series (FTS) method in (Othman and Azahari, 2016). Both methods are verified using two statistical criteria, which are Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). As mentioned in the results, the error values of the suggested method are lesser than the existing method in (Othman and Azahari, 2016). Therefore, this study deduces that the suggested method much precise and effective. Hence, the use of this method can provide a highly accurate determination of intervals to forecast the CPO prices.

ACKNOWLEDGMENTS

Thank you to AP. Dr. Mahmod Bin Othman, my supervisor for his precious guidance and helps throughout this study. Many thank Dr. Rajalingam Sokkalingam, my co-supervisor for a good idea toward this study.

REFERENCES

Ahmad, M. H., Ping, P. Y., & Mahamed N. (2014), "Volatility modelling and forecasting of Malaysian crude palm oil prices," *Appl. Math. Sci.*, **8(124)**, pp. 6159–6169.

Arasu A. & Manku G. S.,(2004) "Approximate counts and quantiles over sliding windows," *Proc. twenty-third ACM SIGMOD-SIGACT-SIGART Symp. Princ. database Syst. - Pod. '04*, no. 2003–72, p. 286.

Asklany S. A., Elhelow K., Youssef I. K., & Abd El-wahab M.,(2011) "Rainfall events prediction using rule-based fuzzy inference system," *Atmos. Res.*, **101(1–2)**, pp. 228–236.

BenYahmed Y., Abu Bakar A., RazakHamdan A., Ahmed A., & Abdullah S. M. S.,(2015) "Adaptive sliding window algorithm for weather data segmentation," *J. Theor. Appl. Inf. Technol.*, **80(2)**, pp. 322–333.

Bingham E., Gionis A., Haiminen N., Hiisila H., Mannila H., & Terzi E.,(2006) "Segmentation and dimensionality reduction," *Proc. Sixth SIAM Int. Conf. Data Min.*, p. 372.

D'Arcy J. A., Collins D. J., Rowland I. J., Padhani A. R., & Leach M. O.,(2002) "Applications of sliding window reconstruction with cartesian sampling for dynamic contrast enhanced MRI," *NMR Biomed.*, **15(2)**, pp. 174–183.

Dani S. & Sharma S,(2013) "Forecasting rainfall of a region by using fuzzy time series," *Asian J. Math. Appl.*, **13(1)**, pp. 1–10.

Datar M., Gionis A., Indyk P., & Motwani R.,(2002) "Maintaining stream statistics over sliding windows:(extended abstract)," *Proc. Thirteen. Annu. {ACM-SIAM} Symp. Discret. algorithms*, vol. 31, no. 6, pp. 635–644.

Huarng K. & Yu T. H.,(2006) "Ratio-Based Lengths of Intervals to Improve Fuzzy Time Series Forecasting," *IEEE Trans. Syst. Man, Cybern. B Cybern.*,**36(2)**, pp. 328–340.

Huarng K.,(2001) "Effective lengths of intervals to improve forecasting in fuzzy time series," *Fuzzy Sets Syst.*,**123(3)**, pp. 387–394.

Jolliffe I. T. & Stephenson D. B.,(2012) *Forecast verification: a practitioner's guide in atmospheric science*. John Wiley & Sons.

Kapoor P. & Bedi S. S.,(2013) "Weather Forecasting Using Sliding Window Algorithm," *ISRN Signal Process.*,**13(10)**, pp. 1–5.

Karia A. A., Bujang I., & Ahmad I.,(2013) "Forecasting Crude Palm Oil Prices using Artificial Intelligence Approach," *Am. J. Oper. Res.*,. **3(13),** pp. 259–267.

Khaliq A. & Ahmad A,(2010) "Fuzzy logic and approximate reasoning,".

Nor S., Azahari F., Othman M., & Saian R.,(2017) "An Enhancement of Sliding Window Algorithm for Rainfall Forecasing," no. 029, pp. 25–30.

Othman M. & Azahari S. N. F.,(2016) "Deseasonalised forecasting model of rainfall distribution using fuzzy time series," *J. Inf. Commun. Technol.*,**15(2)**.

Rahim N. F., Othman M., Sokkalingam R., & Kadir E.,(2018) "Forecasting Crude Palm Oil Prices Using Fuzzy Rule-Based Time Series Method," *IEEE Access*, **6**, pp. 32216–32224.

Rao D. V. N. K., Anusha M., Babu P. N., Sri M. D., & Kumar K. S.,(2015) "Prognostication of Climate Using Sliding Window Algorithm," *Int. J. u- e- Serv. Sci. Technol.*,**8(4)**, pp. 225–232.

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

Vamitha V., Jeyanthi M., Rajaram S., & Revathi T.,(2012) "Temperature Prediction Using Fuzzy Time Series and Multivariate Markov Chain," **2(3)**, pp. 217–230.