Prediction stock price movement using subsethood and weighted subsethood fuzzy time series models

Cite as: AIP Conference Proceedings **2138**, 050018 (2019); https://doi.org/10.1063/1.5121123
Published Online: 21 August 2019

Rosnalini Mansor, Bahtiar Jamili Zaini, and Norhayati Yusof

Prediction Stock Price Movement using Subsethood and Weighted Subsethood Fuzzy Time Series Models

Rosnalini Mansor[1, a)], Bahtiar Jamili Zaini[1, b)], and Norhayati Yusof[1 2,b)]

[1]*School of Quantitative Sciences, UUM College of Arts and Sciences, Universiti Utara Malaysia, 06010 UUM Sintok, Kedah, Malaysia.*
[2]*Quantitative Methods Development & Improvement, UUM College of Arts and Sciences, Universiti Utara Malaysia, 06010 UUM Sintok, Kedah, Malaysia.*

[a)]Corresponding author: rosnalini@uum.edu.my
[b)]bahtiar@uum.edu.my
[c)]norhayati@uum.edu.my

Abstract. Forecasting is the prediction process for the future value. The closing price is usually used to forecast stock price movement in the next period. Predicted stock prices in the investment world become an important thing for stock trading activities. The forecasting process can be the most challenging problems due to difficulty and uncertainty of stock market because stock markets are essentially complex, dynamic, and usually in a nonlinear pattern. One of the novel forecasting methods in this area is fuzzy time series (FTS). This paper proposed stock price movement forecasting using first order and high order weighted subsethood fuzzy time series (WeSuFTS) and subsethood fuzzy time series (SuFTS) methods. A set of secondary data gained from the Kuala Lumpur Stock Exchange (KLSE) website. We chose Malaysian Resources Corp Bhd and we collected the historical data for two months, which is on a day-to-day basis. The performance of four models was analyzed using absolute percentage error (APE), mean square error (MSE), mean absolute percentage error (MAPE) and root mean squared error (RMSE). From the evaluation part of data, the results revealed second order SuFTS is the best model to forecast stock price movement with forecasting error from 0.66% - 6.44% (APE), 2.43% (MAPE), 0.00042 (MSE) and 0.0205 (RMSE).

INTRODUCTION

The stock market has been a center of attraction for the investors for a long period of time. Investors always want to know the expectation of their return on investment in the stock market before they start to invest. The goal is to buy the stock, hold it for a time, and then sell the stock for more than you paid for it. The stock closing price is usually used to forecast stock price movement in the next period. The stock closing price is the price that appears when the stock market closes. The stock closing price is very important because it becomes the reference for the opening price the next day.

Prediction of the stock closing price movement is one of the most challenging problems in the stock market study due to difficulty and uncertainty of stock market [1,2] Stock markets is essentially complex, dynamic, and usually in a nonlinear pattern. The problem is very hard to predict the stock price movements; it is because the stock market volatility needs an accurate forecast model. Stock market indices are highly fluctuating, that is fall the stock price or raising the stock price. In addition, stock market's movement is affected by many macro-economic factors such as political events, firms policies, general economic conditions, commodity price index, bank rate, bank exchange rate, investors' expectations, institutional investors' choice, movements of other stock market and psychology of investors [3].

The 4th Innovation and Analytics Conference & Exhibition (IACE 2019)
AIP Conf. Proc. 2138, 050018-1–050018-6; https://doi.org/10.1063/1.5121123
Published by AIP Publishing. 978-0-7354-1881-3/$30.00

There are many approaches to forecast stock price movement. Linear methods are easy to develop and implement but in reality, the stock movement is in nonlinear patterns [4,5]. Therefore, in recent years, most of the researchers have been concentrating their research work to predict stock price movement by using soft computing methods such as artificial neural network [6,7] and fuzzy set theories [8,9]. However, fuzzy time series (FTS) method has been developed as one of the novel forecasting methods in this area. So far, various FTS methods have been applied successfully to handle stock price forecasting [10-12]. Since this study is focused on applying subsethood fuzzy time series (SuFTS) and weighted subsethood fuzzy time series (WeSuFTS) methods.

Basically, SuFTS method is fuzzy reasoning model embedded in FTS procedure. Then, by using certain weighted calculation from a fuzzy subsethood value in SuFTS, the modified method called WeSuFTS, Previously, WeSuFTS was successfully applied in students enrollment forecasting [13,14] and electricity load demand forecasting [15,16]. Since the previous study conducted WeSuFTS method in first-order WeSuFTS only, there is no attempt for high order WeSuFTS including first and high order SuFTS. Therefore, the objective of this paper is to compare the performance of four models in closing price stock market prediction; first and second order SuFTS and first and second order WeSuFTS models using four error measurements; absolute percentage error (APA), mean squared error (MSE), mean absolute percentage error (MAPE) and root mean squared error (RMSE).

The remainder of this paper is organized as follows: Fuzzy Time Series (FTS) section recalls the basic concepts and definitions that are directly relevant to FTS. The proposed stock price SuFTS and WeSuFTS method are introduced in the next section. and the results are discussed in the results and performance evaluation section and finally, the last section provides conclusion of this study.

FUZZY TIME SERIES

Fuzzy Time Series (FTS) are nonparametric models introduced by Song and Chissom in 1993 [17] based on the fuzzy set theory by Zadeh [18,19]. These methods are easy to implement and very flexible, affording ways to deal with numeric and non-numeric data. Some of the FTS methods produce compact and human readable models of the time series behavior using fuzzy rules which can be used by business experts and researchers. The basic idea of FTS is that historical data are expressed as fuzzy sets and series variation trends are expressed as fuzzy relations. Data are forecasted by fuzzy reasoning while there are not enough historical data or just some imprecise data. FTS approaches have been successfully applied to the data such as stock exchange, temperature, and enrollment which include uncertainty. FTS approaches have found many diversified application areas since it differs from conventional approaches in many respects. The most important is that it does not require the check of theoretical assumptions.

Fuzzy methodologies can provide linguistic-based classification rules. Based on fuzzy methodologies, this approach can create a set of classification rules that capable of classifying the stock price movements into one of the specified groups. In addition, these classification rules can be represented in the form of IF-THEN rules so that it easy to interpret. Therefore, the use of fuzzy methodologies which focuses on the rule generation will help to improve the result on the prediction of the stock movement. Most of the rule generations derive from the membership functions which adapted from the input data. According to Rasmani [20], the weighted subsethood based algorithm can be used for rules generation is because the subsethood values are used as weights over the significance for different conditions available which result in the conclusion. The weight was obtained from the subsethood values then generated the rules for each possible conclusion.

In this study, we focus on fuzzy subsethood and weighted fuzzy subsethood based algorithm to generate the fuzzy logical relationship (FLR) and fuzzy logical relationship group (FLRG) in fuzzy time series. The good thing about subsethood method is this method can show hidden information about the relationship between the fuzzy linguistic input variable and fuzzy output [21].

METHODOLOGY

This section explains the methodology to achieve the objectives of the study. Both type of FTS models; first order and second order FTS used the flow of methodology research as presented in Fig. 1. However, the detail WeSuFTS methodology can refer to Mansor et. al. [13-16]. This paper proposed a methodology FTS model's flow consists of three main phases that include preprocessing data, model development, and model evaluation. The detail steps in each phase are:

Phase 1: Preprocessing data

Step 1: Partitioning data into modelling and evaluation parts.

In this study, the main focus is to predict the future stock price movement for one company listed in Bursa Malaysia. There are more than eight hundred companies listed under Bursa Malaysia. A set of secondary data gained from the Kuala Lumpur Stock Exchange (KLSE) website. We chose Malaysian Resources Corp Bhd and we collected the historical data for two months, which is on a day-to-day basis. The total data in this study is 33 (6 August – 25 September 2018) then 22 data (3 August – 6 September 2018) are prepared for modelling part analysis and 10 data (7 – 25 September 2018) for evaluation part analysis.

Step 2: Data preparation for the FTS models.

Four models were considered in this study:

Model 1: First-order SuFTS
Model 2: First-order WeSuFTS
Model 3: Second-order SuFTS
Model 4: Second-order WeSuFTS.

Lets variable A_t is the closing price at time t, A_{t-1} is previous day closing price and A_{t-2} is a previous two days closing price. Therefore, $A_{t-1} \rightarrow A_t$ is the writing form of first-order FTS model and $A_{t-1}, A_{t-2} \rightarrow A_t$ is the writing form of second order FTS model.

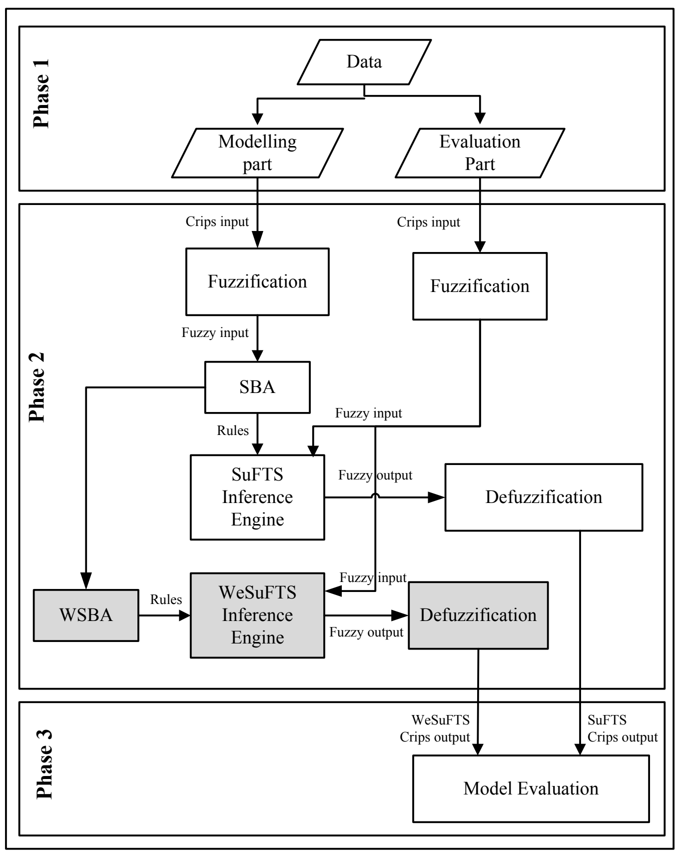

FIGURE 1. SuFTS and WeSuFTS algorithm forecasting stock price movement problem.

Phase 2: Model development.

Step 1: Fuzzifying close price from crisp value into fuzzy value using membership function.

Step 2: (i) Calculate fuzzy subsethood value using subsethood based algorithm (SBA) [20].

The formula calculation is shown in (1). These values describe the relationship between the output linguistic value (A_i) and every input linguistic value (A_j):

$$S(Ai_t, Aj_{t-1}) = \frac{M(Ai_t \cap Aj_{t-1})}{M(Ai_t)} = \frac{\sum_{x \in U} min\{\mu_{Ai_t}(x), \mu_{Aj_{t-1}}(x)\}}{\sum_{x \in U} \mu_{Ai_t}(x)} \quad (1)$$

Step 2; (ii) Calculate weighted fuzzy subsethood value from the value in Step 2(i).
The formula calculation is shown in (2):

$$w(Ai_t, Aj_{t-1}) = \frac{S(Ai_t, Aj_{t-1})}{\max\limits_{j=1,2,3,\dots,6} S(Ai_t, Aj_{t-1})} \quad (2)$$

Step 3: (i) Construct SuFTS model using fuzzy subsethood value in Step 2(i).
The model created is the set of fuzzy reasoning rules. Then we named as SuFTS Inference Engine where we use this proses to make a prediction or forecasting purpose. The min-max operator is used for fuzzy reasoning.

Step 3: (ii) Construct WeSuFTS model using weighted fuzzy subsethood value in Step 2(ii).
Then we named as WeSuFTS Inference Engine where we use this proses to make a prediction or forecasting purpose. The min-max operator is used for fuzzy reasoning.

Step 4: Defuzzifying close price fuzzy value into crisp value using peak point defuzzification method. The defuzzification formula is shown in (3). Where R_i is the calculation value from i rules. M_i is the midpoint or peak point value for i interval in A_t and n is the number of rules ≠ 0.

$$A_t = \frac{\sum_{i=1}^{n} R_i M_i}{\sum_{i=1}^{n} M_i} \quad (3)$$

Phase 3: Model evaluation

Normally, the performance of the forecasting technique is measured with the forecasting error. Although several error measures that are currently available, only four forecasting error measures will be considered in this research because these error measures were the most popular methods among researchers and practitioners. The four forecasting error measures are absolute percentage error (APE), mean squared error (MSE), mean absolute percentage error (MAPE) and root mean squared error (RMSE).

RESULTS AND MODELS PERFORMANCE

This section represents the results from the methodology in the previous section. After completing the fuzzy subsethood value and weighted fuzzy subsethood value calculations, we construct first-order and second-order SuFTS model using subsethood value from equation (1). Also, we construct first-order and second-order WeSuFTS model using weighted subsethood value from equation (2). All the models are shown in Fig. 2,3,4 and 5. After that, we forecast next day close price in data evaluation part using these models by equation (3). The results are shown in Table 1.

Rule 1:	-
Rule 2:	*If A_{t-1} is ($0.56A2_{t-1}$ or $0.56A3_{t-1}$) then A_t is $A2_t$*
Rule 3:	*If A_{t-1} is ($0.0377A1_{t-1}$ or $0.2264A2_{t-1}$ or $0.6604A3_{t-1}$ or $0.3019A4_{t-1}$) then A_t is $A3_t$*
Rule 4:	*If A_{t-1} is ($0.3462A3_{t-1}$ or $0.4615A4_{t-1}$ or $0.3462A5_{t-1}$) then A_t is $A4_t$*
Rule 5:	*If A_{t-1} is ($0.0152A3_{t-1}$ or $0.3182A4_{t-1}$ or $0.8182A5_{t-1}$ or $0.1061A6_{t-1}$) then A_t is $A5_t$*
Rule 6:	-

FIGURE 2. 1st order SuFTS close price model

Rule 1:	-
Rule 2:	*If A_{t-1} is ($A2_{t-1}$ or $A3_{t-1}$) then A_t is $A2_t$*
Rule 3:	*If A_{t-1} is (0.0571$A1_{t-1}$ or 0.3429$A2_{t-1}$ or $A3_{t-1}$ or 0.4571$A4_{t-1}$) then A_t is $A3_t$*
Rule 4:	*If A_{t-1} is (0.75$A3_{t-1}$ or $A4_{t-1}$ or 0.75$A5_{t-1}$) then A_t is $A4_t$*
Rule 5:	*If A_{t-1} is (0.0185$A3_{t-1}$ or 0.3889$A4_{t-1}$ or $A5_{t-1}$ or 0.1296$A6_{t-1}$) then A_t is $A5_t$*
Rule 6:	-

FIGURE 3. 1st order WeSuFTS close price model

Rule 1:	-
Rule 2:	*If A_{t-1} is (0.56$A2_{t-1}$ or 0.56$A3_{t-1}$) and A_{t-2} is (0.24$A2_{t-2}$ or 0.88$A3_{t-2}$) then A_t is $A2_t$*
Rule 3:	*If A_{t-1} is (0.0377$A1_{t-1}$ or 0.2264$A2_{t-1}$ or 0.6604$A3_{t-1}$ or 0.3019$A4_{t-1}$) and A_{t-2} is (0.0377$A1_{t-2}$ or 0.283$A2_{t-2}$ or 0.434$A3_{t-2}$ or 0.4151$A4_{t-2}$ or 0.0377$A5_{t-2}$) then A_t is $A3_t$*
Rule 4:	*If A_{t-1} is (0.3462$A3_{t-1}$ or 0.4615$A4_{t-1}$ or 0.3462$A5_{t-1}$) and A_{t-2} is (0.2692$A3_{t-2}$ or 0.3846$A4_{t-2}$ or 0.5$A5_{t-2}$) then A_t is $A4_t$*
Rule 5:	*If A_{t-1} is (0.0152$A3_{t-1}$ or 0.3182$A4_{t-1}$ or 0.8182$A5_{t-1}$ or 0.1061$A6_{t-1}$) and A_{t-2} is (0.1212$A3_{t-2}$ or 0.2576$A4_{t-2}$ or 0.697$A5_{t-2}$ or 0.1061$A6_{t-2}$) then A_t is $A5_t$*
Rule 6:	-

FIGURE 4. 2nd order SuFTS close price model

Rule 1:	-
Rule 2:	*If A_{t-1} is ($A2_{t-1}$ or $A3_{t-1}$) and A_{t-2} is (0.2727$A2_{t-2}$ or $A3_{t-2}$) then A_t is $A2_t$*
Rule 3:	*If A_{t-1} is (0.0571$A1_{t-1}$ or 0.3429$A2_{t-1}$ or $A3_{t-1}$ or 0.4571$A4_{t-1}$) and A_{t-2} is (0.087$A1_{t-2}$ or 0.6522$A2_{t-2}$ or $A3_{t-2}$ or 0.9565$A4_{t-2}$ or 0.087$A5_{t-2}$) then A_t is $A3_t$*
Rule 4:	*If A_{t-1} is (0.75$A3_{t-1}$ or $A4_{t-1}$ or 0.75$A5_{t-1}$) and A_{t-2} is (0.5385$A3_{t-2}$ or 0.7692$A4_{t-2}$ or $A5_{t-2}$) then A_t is $A4_t$*
Rule 5:	*If A_{t-1} is (0.0185$A3_{t-1}$ or 0.3889$A4_{t-1}$ or $A5_{t-1}$ or 0.1296$A6_{t-1}$) and A_{t-2} is (0.1739$A3_{t-2}$ or 0.3696$A4_{t-2}$ or $A5_{t-2}$ or 0.1522$A6_{t-2}$ then A_t is $A5_t$*
Rule 6:	-

FIGURE 5. 2nd order WeSuFTS close price model

TABLE 1. Models performance results

time(t)	Actual value	1st order model				2nd order model			
		WeSuFTS model		SuFTS model		WeSuFTS model		SuFTS model	
		Forecasted value	APE	Forecasted value	APE	Forecasted value	APE	Forecasted value	APE
23	0.720	0.7420	3.0604	0.7300	1.3928	0.7324	1.7158	0.7300	1.3928
24	0.690	0.7486	8.4888	0.7376	6.8920	0.7429	7.6645	0.7344	6.4373
25	0.700	0.7275	3.9343	0.7144	2.0564	0.7203	2.9058	0.7165	2.3592
26	0.705	0.7248	2.8150	0.7144	1.3326	0.7306	3.6253	0.7271	3.1284
27	0.715	0.7300	2.0953	0.7187	0.5113	0.7265	1.6146	0.7230	1.1131
28	0.725	0.7420	2.3496	0.7300	0.6936	0.7351	1.3904	0.7298	0.6630
29	0.710	0.7545	6.2619	0.7442	4.8173	0.7435	4.7229	0.7347	3.4840
30	0.730	0.7356	0.7627	0.7234	0.8996	0.7315	0.2111	0.7237	0.8625
31	0.720	0.7531	4.5928	0.7442	3.3615	0.7449	3.4529	0.7375	2.4366
32	0.720	0.7486	3.9685	0.7376	2.4382	0.7449	3.4529	0.7375	2.4294
MAPE		3.8329		2.4395		3.0756		2.4306	
MSE		0.00095		0.00048		0.00066		0.00042	
RMSE		0.0308		0.0219		0.0257		0.0205	

The performance of the SuFTS and WeSuFTS models with a different order or a different number of lags were measured by Absolute Percentage Error (APE) Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). Table 1 shows the forecasting closing price stock market, APE, MAPE, MSE and RMSE for each model. The performance comparison results show that the second-order SuFTS has APE values range from 0.66% – 6.44%, MAPE value is 2.43%, MSE value is 0.004 and RMSE value is 0.0205. Hence, these results show that second-order SuFTS is outperform compared to the other three models in this study in forecasting close price movement in the stock market.

CONCLUSION

Fuzzy Time Series (FTS) is becoming quite popular since it was introduced by Chisom and Chen in 1993. Many related kinds of research are conducted in student enrolment and stock market problems. Two of the FTS model category is the first order and high order FTS model. This study proposed first and second order SuFTS and WeSuFTS close price stock market models. The proposed algorithm in this study employs subsethood fuzzy time series (SuFTS) and weighted subsethood fuzzy time series (WeSuFTS). SuFTS and WeSuFTS are the fuzzy reasoning based time series forecasting using fuzzy subsethood model and weighted fuzzy subsethood model.

This paper gives a step by step of SuFTS and WeSuFTS and it is a still new method in FTS since it presents a forecasting model based on fuzzy reasoning rule model based on fuzzy subsethood value. The proposed method uses the min-max operator for fuzzy reasoning and peak-point defuzzification which make the method simple and less complex.

ACKNOWLEDGMENT

The authors are grateful for the financial support received from Universiti Utara Malaysia under the Research Generating University Grant (S/O Code: 13878).

REFERENCES

1. H. Chulia, M. Guillen and J. M. Uribe, Int. Rev. Econ. Financ. **48**, pp. 18-33 (2017).
2. R. Connolly, C. Stivers and L. Sun, J. Financ. Quant. Anal. **40**, pp. 161-194 (2005).
3. A.M Adam and G. Tweneboah, SSRN Electro. J. 1-26 (2008).
4. F. Ye, L. Zhang, D. Zhang, H. Fujita, and Z. Gong, Inf. Sci. (Ny)., **41**, pp. 367–368 (2016).
5. W.-J. Chuang, L.-Y. Ou-Yang, and W.-C. Lo, Analele Stiint. Ale Univ. Alexandru Ioan Cuza" Din Iasi-Stiinte Econ. **56**, pp. 621-634 (2009).
6. Y.-G. Song, Y.-L. Zhou, and R.-J. Han, J. Differ. Equations Appl. **00**, pp. 1-13 (2018).
7. Ö. İcan and T.B. Çelik, Int. J. Econ. Financ. **9**, pp. 100 (2017).
8. H. Guan, Z. Dai, A. Zhao, and J. He, PLoS One **13**, pp. 1-15 (2018).
9. X. Pang, Y. Zhou, P. Wang, W. Lin, and V. Chang, J. Supercomput. pp. 1-21 (2018).
10. Y. Wang, Y. Lei, X. Fan, and Yi Wang, Math. Probl. Eng. **2016**, pp. 1-12 (2016).
11. B.P. Joshi and S. Kumar, Int. J. Model. Simulation, Sci. Comput. **04**, pp. 1-12 (2013).
12. S.M. Chen and C.D. Chen, IEEE Trans. Fuzzy Syst. **19**, pp. 1-12 (2011).
13. R. Mansor, M.M. Kasim, and M. Othman, in *SKSM 25. 2018* (AIP Publishing, Pahang, 2017), p. 040017-1 - 040017-7.
14. R. Mansor, M. Othman, and M.M. Kasim, Adv. Sci. Lett. **23**, pp. 9094-9097 (2017).
15. R. Mansor, M.M. Kasim, and M. Othman, in *Int. Conf. Appl. Sci. Technol. 2016* (AIP Publishing, Langkawi, 2016), pp. 20061.
16. R. Mansor, M.M. Kasim, and M. Othman, J. Telecommun. Electron. Comput. Eng. **8**, pp. 97-102 (2016).
17. Q. Song and B.S. Chissom, Fuzzy Sets Syst. **54**, pp.1-9 (1993).
18. L. Zadeh, Inf. Science **8**, pp. 199-249 (1975).
19. L. Zadeh, Inf. Control **8**, pp. 338-353 (1965).
20. K.A. Rasmani, in *IEEE Int. Conf. Fuzzy Syst.* (Budapest, 2004), pp. 1679–1684.
21. S.-M. Chen, S.-H. Lee, and C.-H. Lee, Appl. Artif. Intell. An Int. J. **15**, pp. 645-664 (2001).