Volume 2 No. 1 ISSN 2079-8407
Journal of Emerging Trends in Computing and Information Sciences

http://www.cisjournal.org

Fourth-order Fuzzy Time Series Based on Multi-Period Adaptation Models for Kuala Lumpur Composite Index Forecasting

Lazim Abdullah
Department of Mathematics
University of Malaysia Terengganu
21030 Kuala Terengganu, Malaysia
lazim@umt.edu.my

Toh Lian Fan
Department of Mathematics
University of Malaysia Terengganu
21030 Kuala Terengganu, Malaysia
liangfan87@hotmail.com

ABSTRACT

Forecasting accuracy is one of the most critical issues in fuzzy time-series models. A combination model of higher order fuzzy time series with adaption period was proposed by Chen et al as one of the mechanisms in improving forecasting accuracy. First, second, third and fourth order fuzzy times series based on multi-period adaptation models were successfully tested with stock exchange indexes in Taiwan and Hong Kong. However feasibility of the fourth order model especially the effect of multi-period adaptations to local indexes remains unknown. Therefore the present paper tests the fourth-order based on one, two, three and four-period adaptation model with datasets of Kuala Lumpur Composite Index. With the fourth-order of the adaptation model, it is found that the forecasted index based on two-period adaptation performed better than the other adaptation periods. The fourth-order based on two-period adaption seems perfectly worked with datasets of Kuala Lumpur Composite Index.

Keywords: *fourth-order, fuzzy time series, composite index, multi-period adaptation, forecasting.*

I. INTRODUCTION

It is generally known that there are two approaches that can be used to predict the stock market prices. The approaches are fundamental analysis and technical analysis. Fundamental analysis is a technique that a security's value is determined by study on everything that can affect the security's value, including macroeconomic factors such as overall economy and industry conditions and also company-specific factors like financial condition and management. On the others hand, technical analysis is a security analysis that use the study of past market data, primarily price and volume for forecasting the future direction of prices. In economic forecasting, especially in the area of stock index forecasting, the most common method that has been used is time-series. Traditional time-series methods such as autoregressive models, moving average models, and Autoregressive Integrated Moving Average model can be applied in forecasting economic problem, but there are fail to forecast the problems with linguistic historical data. Furthermore, traditional time-series methods require more historical data and the data must be normally distributed to obtain a better forecasting performance. However, there are often not sufficient data in all kinds of time-series model, and linguistic expressions are often used to describe the daily observations. Since [1] proposed the original model, fuzzy time-series model has been employed to deal with various domains forecasting problems, such as university enrollment forecasting, temperature forecasting and stock index forecasting.

Reference [2], [3] have presented a simple fuzzy composition method and high-order fuzzy time-series in order to improve the forecasting accuracy rate of the fuzzy time-series. In area of stock price forecasting, [4] proposed heuristic models to improve the forecasting performance by integrating problem-specific heuristic knowledge with Chen's model, which was proposed to forecast university enrollment. In the same year, [5] has found that the length of linguistic intervals for the universe of discourse may affect forecasting results and proposed two linguistic interval partitioning approaches, distribution-based and average-based length, to approach this issue. Although the forecasting performance of Huarng's model demonstrates excellence in forecasting, it creates too many linguistic values to be identified by analysts.

In year 2005, [6] proposed another model which used weighted method to tackle the two issues that is recurrence and weighting in fuzzy time-series forecasting. Reference [7] had proposed a trend-weighted model to echo Yu's research. Reference [8] presented a two-factor high-order fuzzy time-series for temperature prediction and Taiwan futures exchange forecasting. It is noted that the forecasting accuracy rates of these methods mainly depend on their universe of discourse and the length of intervals.

In time-series methods, the process determining the length of linguistic intervals and mining Fuzzy Logical

Relationships (FLR) from time-series are critical factor to influence forecasting accuracy. Therefore, some advanced algorithms such as genetic algorithms and neural networks were applied to improve this process. Reference [9] had presented a new method which was using the high-order fuzzy time-series and genetic algorithms for forecasting enrollments. Although advanced algorithms perform well in forecasting, they only deal with non-linear relationships in time-series. In stock markets, investors usually make their investment decisions based on recent stock information such as latest market news, stock technical indicators, or price fluctuations.

The major problem in forecasting is forecasting accuracy. To reconcile this problem, [10] developed high-order fuzzy time-series and modify the forecasts with multi-period adaptation model which is derived from adaptive expectation model to enhance forecasting accuracy. They tested the model to experimental datasets at Taiwan Stock Exchange Capitalization Weighted Stock Index and Hang Seng Index from first-order to fourth-order based on one, two, three and four–adaptation period. The performances of the models were associated not only with order of the model but also number of adaptation period. Number of adaptation periods from one to four opens a new search in forecasting accuracy especially at the highest order. Furthermore, the model remains open to be tested with other composite indexes. Therefore the aim of this paper is to test the fourth-order of fuzzy time series with multi-period adaptation model from one to four to the datasets of Kuala Lumpur Composite Index (KLCI). This paper is structured as follows. Definitions of fuzzy time series and its affiliates are presented in Section II. Section III explains the high-order of fuzzy time series. Computational steps leading to the forecasting results specifically for KLCI datasets and performances are laid in Section IV. Results of the forecasting indexes and performances are given in Section V. The paper finally ends with a short conclusion in Section VI.

II. PRELIMINARIES

Definitions and procedures of the fuzzy time-series presented by [1] are described as follows:

Definition 1: Let Y(t) (t= 0,1, 2,…) is a subset of real numbers, Y(t) be the universe of discourse by which fuzzy sets fj(t) are defined. If F(t) is a collection of f1(t), f2(t)……, then F(t) is called a fuzzy time-series defined on Y(t).

Definition 2: If there exists a fuzzy logical relationship R(t-1, t) such that F(t) = F(t-1)$_{o}$R(t-1,t) where "$\circ$" represents the max–min composition operator, F(t-1) and F(t) are fuzzy sets, then F(t) is said to be caused by F(t-1). The logical relationship between F(t) and F(t-1) can be represented as: F(t-1)→F(t).

Definition 3: Suppose F(t) is caused by F(t-1) only, and F(t) =F(t-1)$\times$R(t-1,t). For any t, if R(t-1,t) is independent of t, then F(t) is named a time-invariant fuzzy time-series, otherwise a time-variant fuzzy time-series.

Definition 4: Let **F (t-1) = Ai** and **F(t)=Aj**. The relationship between two consecutive observations, F (t) and F (t-1), referred to as a fuzzy logical relationship (FLR), can be denoted by Ai→Aj, where Ai is called the Left-Hand Side (LHS) and Aj the Right-Hand Side (RHS) of the FLR.

Definition 5: All fuzzy logical relationships in the training dataset can be further grouped together into different fuzzy logical relationship groups according to the same Left-Hand Sides of the fuzzy logical relationship. For example, there are two fuzzy logical relationships with the same Left-Hand Side (Ai): Ai →Aj1 and Ai→Aj2. These two fuzzy logical relationships can be grouped into a fuzzy logical relationship group which is Ai →Aj1,Aj2. .

Definition 6: Assume that F(t) is a fuzzy time-series and F(t) is caused by F(t-1),F(t-2) ,….F(t-n), then the nth-order fuzzy logical relationship can be represented as follows: F(t-1), F(t-2),….F(t-n)→F(t) where F(t-1) F(t-2), …..F(t-n) are called the antecedent and F(t) is called the consequent of the nth-order fuzzy time-series forecasting model, where n ≥ 2.

III. HIGH-ORDER FUZZY TIME-SERIES

In the research of fuzzy time-series, most critical factors to influence the forecasting accuracy are the process determining the length of linguistic intervals and mining fuzzy logical relationships. In stock market, investors usually make their short-term investment decisions according to recent stock information such as the late market news, price fluctuations, or stock technical indicators.

Based on these assumptions, the previous time periods (t-1, t-2, t-3,t-4) of stock index should be taken as forecasting attributes (first-order, second-order, third-order and fourth-order) for the future stock index. So, a thoughtful fuzzy time-series model should consider two price patterns together in forecasting processes: (1) nonlinear high-order fuzzy logical relationship data; and (2) linear relationships between recent periods of stock prices. Therefore, in Chen's model, a high-order fuzzy time-series model is proposed to implement this concept.

Assume that F(t) is a fuzzy time-series and F(t) is caused by F(t-1), F(t-2),…, F(t-n), then the nth-order fuzzy logical relationship can be defined as in Equation (3.1)

$$\mathbf{F(t-1), F(t-2),\dots, F(t-n) \rightarrow F(t)} \qquad (3.1)$$

where F(t-1), F(t-2),…, F(t-n) are called the antecedent and F(t) is called the consequent of the nth-order fuzzy time-series forecasting model, where n ≥ 2.

According to the previous research, [10] and [11] had employed the multi-period adaptation equation which derived from the adaptive expectation model [12] to produce a conclusive forecasting value and to enhance forecasting accuracy. In time series research, adaptive expectation model is a reasonable forecast model to represent the prediction approach of stock investors for the future stock price, where the forecast for the future stock price is generated by the last period of stock price and the correction of forecasting error.

The adaptive expectation model is defined as

$$P(t) = P(t-1) + h_0 * \varepsilon(t-1) \quad \textbf{(3.2)}$$

where P(t) is the price at time t; P(t-1) is the price on time t-1; $\varepsilon(t-1)$ is the defuzzification error at time t-1 and $h_0$is a weighted parameter.

If the future stock price is influenced by recent i periods of stock prices or stock investors make their decisions based on recent i periods of stock prices, then a multi-period adaptation equation can derived from the extending formula of the adaptive expectation model based on multiple periods of error corrections. Equation (3.3) defined the multi-period adaptation model.

$$\text{Adaptive\_forecast}(t+1) = P(t) + \sum_{i=1}^{k} h_i * [\text{Forecast}(t+1) - P(t)] \quad \textbf{(3.3)}$$

where Adaptive forecast (t +1) is the conclusive forecast for the future stock price; P(t) is the actual stock price on time t; Forecast (t +1) is the initial forecasting value from equation (3.3) and h_i is a linear parameter.. The linear parameter, h_i range from -1 to 1 but 0 with the stepped value, 0.001.

IV. COMPUTATION

The seven steps for the high-order fuzzy time-series based on multi-period adaptation model to forecast the KLCI is executed as follows. The forecasting performances with of the fourth-order based on multiple-adaptation periods, from one to four periods, are also computed.

Step 1. Define Universe of Discourse, U

In this step, the universe of discourse is defined as $U = [D_{min} - D_1, D_{max} + D_2]$ to preserve a variation space to ensure the future stock index are encompassed in U, where D_1 and D_2 are two real positive numbers and D_{min} and D_{max} are the minimum and maximum values of the training dataset. The maximum and minimum values of the KLCI stock index in the training dataset are 1516.22 and 829.41, respectively. Then the universe of discourse can be defined as $U = [800, 1600]$ where D_1 is 29.41 and D_2 is 83.78.

Step 2. Partition Universe of Discourse

Partition the universe of discourse is required in fuzzy time series to produce several equal lengths of linguistic intervals for the training datasets. The suggestion from [13] reckoned that the universe of discourse should be partitioned into seven linguistic values, or seven plus or minus two. Hence, partitioning into eight equal length intervals is opted and it is shown in Table I.

TABLE I THE EIGHT LINGUISTIC INTERVALS FOR KLCI

Linguistic Intervals	Interval Range	Midpoint
u_2	[800-900)	850
u_2	[900-1000)	950
u_3	[1000-1100)	1050
u_4	[1100-1200)	1150
u_5	[1200-1300)	1250
u_6	[1300-1400)	1350
u_7	[1400-1500)	1450
u_8	[800-900)	850

Step 3. Establish Fuzzy Set for Each Observation

$L_1, L_2, \ldots, L_k$ be fuzzy set and the fuzzy set, $L_1, L_2, \ldots, L_k$, based on the universe of discourse U are defined by

$$L_1 = 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + 0/u_5 + 0/u_6 + 0/u_7 + 0/u_8$$
$$L_2 = 0.5/u_1 + 1/u_2 + 0.5/u_3 + 0/u_4 + 0/u_5 + 0/u_6 + 0/u_7 + 0/u_8$$
$$L_3 = 0/u_1 + 0.5/u_2 + 1/u_3 + 0.5/u_4 + 0/u_5 + 0/u_6 + 0/u_7 + 0/u_8$$
$$L_4 = 0/u_1 + 0/u_2 + 0.5/u_3 + 1/u_4 + 0.5/u_5 + 0/u_6 + 0/u_7 + 0/u_8$$
$$L_5 = 1/u_1 + 0.5/u_2 + 0/u_3 + 0.5/u_4 + 1/u_5 + 0.5/u_6 + 0/u_7 + 0/u_8$$
$$L_6 = 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + 0.5/u_5 + 1/u_6 + 0.5/u_7 + 0/u_8$$
$$L_7 = 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + 0/u_5 + 0.5/u_6 + 1/u_7 + 0.5/u_8$$
$$L_8 = 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + 0/u_5 + 0/u_6 + 0.5/u_7 + 1/u_8$$

The eight fuzzy linguistic values can be defined as follows: L_1 = (very low stock index), L_2 = (median low stock index), L_3 = (low stock index), L_4 = (little low stock index), L_5 = (little high stock index), L_6 = (high stock index), L_7 = (median high stock index) and L_8 (very high stock index). Convert each stock price in training datasets to corresponding linguistic values. For example, at time t = 1,

Volume 2 No. 1 ISSN 2079-8407
Journal of Emerging Trends in Computing and Information Sciences

http://www.cisjournal.org

the KLCI value is assigned to L_4. It is because its index value is equal to 1118.18 which is fall in the interval range of [1100-1200).

Step 4. Establish i th-order Fuzzy Logical Relationships (FLR)

The fuzzy logical relationships (FLR) are generated based on the fuzzified observation. FLRs for the training datasets are established based on linguistic values in Step 3. First-order fuzzy logical relationship is composed of two consecutive linguistic values, the FLR $(L_i \rightarrow L_j)$ is established by $L_i(t-1)$ and $L_j(t)$. For example,

$$
\begin{aligned}
&L_6\ (t=211) \rightarrow L_7\ (t=212)\\
&L_7\ (t=212) \rightarrow L_7\ (t=213)\\
&L_7\ (t=213) \rightarrow L_7\ (t=214)\\
&L_7\ (t=214) \rightarrow L_7\ (t=215)\\
&L_7\ (t=215) \rightarrow L_7\ (t=216)\\
&L_6\ (t=216) \rightarrow L_6\ (t=217)\\
&L_6\ (t=217) \rightarrow L_6\ (t=218)\\
&L_6\ (t=218) \rightarrow L_7\ (t=219)\\
&\qquad\qquad \vdots
\end{aligned}
$$

Second-order fuzzy logical relationship is composed of three consecutive linguistic values, the FLR $(L_i, L_j \rightarrow L_k)$ is established by $L_i(t-2)$, $L_j(t-1)$ and $L_k(t)$. Third-order fuzzy logical relationship is composed of four consecutive linguistic values, the FLR $(L_i, L_j, L_k \rightarrow L_l)$ is established by $L_i(t-3)$, $L_j(t-2)$, $L_k(t-1)$ and $L_l(t)$.Fourth-order fuzzy logical relationship is composed of five consecutive linguistic values, the FLR $(L_i, L_j, L_k, L_l \rightarrow L_m)$ is established by $L_i(t-4)$, $L_j(t-3)$, $L_k(t-2)$, $L_l(t-1)$ and $L_m(t)$.

Step 5. Establish FLR Groups and Produce Frequency Weight Matrix of Each FLR

Fuzzy logical relationships with the same left hand side (LHS) linguistic values can be grouped into one FLR group.

In the frequency-weighted method, each FLR weight is determined by its occurrence frequency. The corresponding weights for $L_1, L_2, \ldots, L_8$, , say $W_1, W_2, \ldots, W_8$, , are specified. However, before forming the weight matrix with these $W_1, W_2, \ldots, W_8$, the weight matrix, $W(t) = [W_1', W_2', \cdots, W_8']$ should satisfy condition

$$\sum_{i=1}^{k} W_i' = 1$$

Step 6. Establish i th-order Fuzzy Logical Relationships (FLR)

In this step, two matrixes are used to generate initial forecasts. This process is called defuzzification. L_{df} is the defuzzified matrix which is composed of each midpoint of the linguistic interval.

Step 7. Establish i th-order Fuzzy Logical Relationships (FLR)

Lastly, the initial forecast results which computed in previous step are substitute into multi-period adaptation model.

Performance of the forecasting results is examined using Root Mean Squared Error (RMSE). The error formula is given as

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} [actual(t) - forecast(t)]^2}{n}}$$

V. Results

Datasets of KLCI yield forecasting results and performances via the high-order fuzzy time series based on multiple-adaptation models. Govern by limitation of this paper, the fourth-order fuzzy time series with one, two, three and four-adaptation period models for KLCI forecasting can be depicted in Figure 1

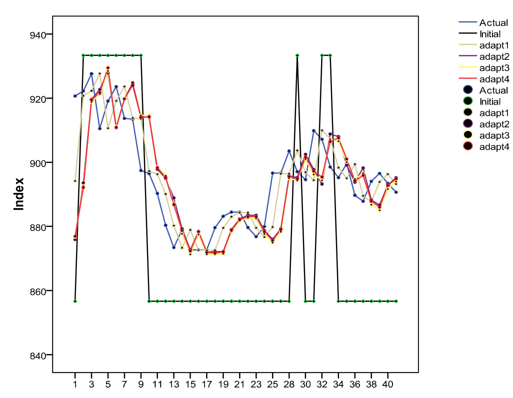

Figure I Forecasting Lines For The Model

As can be seen from the Fig I, the effect of adaptation periods largely contributed to the forecasting performances. The lines of one, two, three and four-adaptation period are quite close to the actual line.

The performance of each adaptation period can be further examined using error analysis. Using the RMSE formula, performances of the model can be summarized in Table II.

TABLE II ERROR ANALYSIS FOR INITIAL AND ADAPTATION PERIOD FORECASTING

	Initial	1 adaptation period	2 adaptation period	3 adaptation period	4 adaptation period
RMSE	41.24	9.31	9.01	9.26	9.26

It can be observed that the fourth-order model produced the best forecasted index at second-adaptation period. It shows that the investment patterns at Kuala Lumpur Stock Exchange are short-term.

VI. Short Conclusion

According to the study of [14], they suggest that combining of several models can be an effective way to improve forecasting performance and also reduces the risk in forecasting failure. In this paper, combination of fourth-order fuzzy time series with the multi-period adaptation model has been employed to forecast the KLCI stock index. The results show that the forecasting model that combines with the multi-period adaptation model produce lower RMSE compared with the model without adaptation period. It shows the multi-period adaptation models can effectively improve the forecasting performance.

References

[1] Q.Song, and B.S.Chissom,. Fuzzy time-series and its models, Fuzzy Sets and Systems 54, 1993, 269-277.

[2] S.M.Chen,. Forecasting enrollments based on fuzzy time series, Fuzzy Set and Syatems, 81, 1996, 311-319.

[3] S.M Chen,. Forecasting enrollments based on high-order fuzzy time series, Cybernetics and Systems 33, 2002, 1-16

[4] K.H.Huarng,. Heuristic models of fuzzy time series for forecasting, Fuzzy Sets and System 123, 2001 , 369-386.

[5] K.H. Huarng,. Effective lengths of interval to improve forecasting in fuzzy time series, Fuzzy Sets and System 123, 2001, 155-162

[6] H.K.Yu,.Weighted fuzzy time series models for TAIEX forecasting, Physica A 349, 2005, 609-624

[7] C.H.Cheng, , T.L.Chen, , and C.H.Chiang, Trend-weighted fuzzy time-series model for TAIEX forecasting, Lecture Notes in Computer Science 4234, 2006, 469-477.

[8] L.W. Lee, L.H Wang,., S.M.Chen, and Y.H. Leu.. Handling forecasting problems based on two-factor high-order fuzzy time-series. IEEE Transactions on Fuzzy Systems, 14(3), 2006, 468-477.

[9] S.M. Chen, and N.Y.Chung,. Forecasting enrollments using high-order fuzzy time series and genetic algorithms, International Journal of Intelligent Systems 21, 2006, 485-501.

[10] T.L.Chen, C.H.Cheng, and H.J. Teoh,. High-order fuzzy time-series on multi-period adaptation model for forecasting stock markets, Physical A 38, 2008, 876-888.

[11] T.L Chen, C.H.Cheng, , and H.J.Teoh. Fuzzy time-series based on Fibonacci sequence for stock price forecasting, Physical A 380, 2007, 377-390.

[12] J.Kmenta. Elements of Econometrics, MacMillan: University of Michigan Press. 1986.

[13] G.A. Miler.The magical number seven, plus or minus two:
some limits on our capacity of processing information, Psychol. Rev. 63, 1956, 81-97

[14] M,Hibon, and T. Evgeniou,. To combine or not to combine: selecting among forecasts and their combinations, International Journal of Forecasting 21,2005, 253-280.