Fuzzy time series Markov Chain and Fuzzy time series Chen & Hsu for forecasting

by Susilo Hariyanto

Journal of Physics: Conference Series **1943** (2021) 012128 doi:10.1088/1742-6596/1943/1/012128

Fuzzy time series Markov Chain and Fuzzy time series Chen & Hsu for forecasting

Zaenurrohman[1], S Hariyanto[2], T Udjiani[2]

[1]Magister Program of Mathematics, Departement of Mathematics, Faculty of Science and Mathematics, Diponegoro University
[2]Departement of Mathematics, Faculty of Science and Mathematics, Diponegoro University

Corresponding author: zaenurrohman@students.undip.ac.id

Abtract. Fuzzy time series theory is a concept of artificial intelligence that can use to conduct forecasting technique.. This paper discusses the fuzzy logic concept to develop the base of the fuzzy time series with time invariant and time variant methods. There are several methods of fuzzy time series, including Markov Chain method and Chen and Hsu method. The Markov Chain method combines between the fuzzy time series and the Markov Chain. This merger goals to finest opportunity of the use matrix probability transitions. Chen and Hsu method is based on the historical data difference in conducting forecasting. By using Markov Chain and Chen & Hsu methods, it may achieve forecasting outcomes with a low mistakes rate. To clarify each technique and for comparison further, it is given an example of the relevant issue to be resolved by both methods. The consequences acquired can be compared, so it can be concluded which method is better.

1. Introduction

Forecasting is an activity estimating what happens in the future for a fantastically lenghthy time. Forecasting is an essential trouble spanning many fields inclusive of enterprise and industry, government, economics, environmental science, medical, social, politics, and finance[1]. In economics, forecasting is an crucial a part of the business enterprise whilst making control decisions. In forecasting is needed to a narrowest feasible error.

One of the techniques currently developed to perform forecasting is the Fuzzied Time Series Technique which is classified into artificially intelligent concepts or artificial intelligence concepts that can help to perform forecasting techniques. This technique turned into first proposed by Song and Chissom who used the fuzzy logic concept to develop the basis of the fuzzy time series using the time invariant and time variant methods used to carry out forecasting[2]. Some of the fuzzy time series methods had been developed including the Markov method [3], Chen[4], Chen and Hsu[5], the weighted method [6], the multiple-attribute fuzzy time series method[7], the percentage change method [8], and the markov chain method[9]. To date there has been no really precise method of forecasting the future value of the share price.

Since there may be no truly particular approach yet, this paper will examine two methods namely Markov Chain and Chen & Hsu to be compared which method is better. The Markov Chain and Chen & Hsu strategies had been selected due to the fact they have a got an excessive diploma of accuracy.

Journal of Physics: Conference Series **1943** (2021) 012128 doi:10.1088/1742-6596/1943/1/012128

2. Discussion

2.1. *Fuzzy time series*

Suppose U is a set of the universe, U = $\{u_1, u_2, \ldots, u_n\}$, then the fuzzy set A of U is defined as follows:

$$A = \frac{f_A(u_1)}{u_1} + \frac{f_A(u_2)}{u_1} + \frac{f_A(u_3)}{u_1} + \cdots + \frac{f_A(u_n)}{u_1} \tag{1}$$

where f_A is a function member of the fuzzy set A, $f_A : U \rightarrow [0,1]$, $f_A(u_i)$ denotes the grade of club of u_i in the fuzzy set A, and $1 \leq i \leq n$ [10].

Definition 2.1.1. [5]

Suppose $(t)(t = \cdots, 0,1,2, \ldots)$, a subset of **R**, becomes a universe on a set of fuzzy $f_i(t)$ (i = 1,2, ...). Then, $F(t)$ is referred to as fuzzy time series on $(t)(t = \cdots, 0,1,2, \ldots)$.

From the above definition, it may be seen that F(t) can be considered a linguistic variable and $f_i(t)(i = 1,2, \ldots)$ is seen as the linguistic value of *F(t)*, in which $f_i(t)(i = 1,2, \ldots)$ is a representation of a fuzzy set. We can see that *F(t)* is a function at t time.

Definition 2.1.2. [4]

If *F(t)* is brought on by *F(t - 1)*, i.e., $F(t-1) \rightarrow F(t)$, it can be stated as follows:

$$F(t) = F(t-1) \times R(t, t-1) \tag{2}$$

where $R(t, t-1)$ is the fuzzy relationship between $F(t-1)$ and $F(t)$, and $F(t) = F(t-1) \cdot R(t, t-1)$ is called the first forecasting model on $F(t)$. The relation $F(t-1) \rightarrow F(t)$ is called a *fuzzy logical relationship*, whilst $F(t-1)$ is the current state and $F(t)$ is the subsequent state.

Definition 2.1.3[9]

Suppose among $F(t) = A_i$, is because of $F(t-1) = A_j$, then the fuzzy logical relationship is described as $A_i \rightarrow A_j$.

If there are fuzzy logical relationships acquired from state A_2, then a change is made to any state A_j, $j = 1,2, \ldots, n$, as $A_2 \rightarrow A_3, A_2 \rightarrow A_2, \ldots, A_2 \rightarrow A_1$; *hence*, the fuzzy logical relationships are grouped right into a fuzzy logical relationships group as

$$A_2 \rightarrow A_1, A_2, A_3$$

In general the steps of the fuzzy time series model include: (1) determining the universe of talks, where the fuzzy set will be defined (2) dividing the universe set into multiple intervals of the same length (3) defining the fuzzy set A (4) fuzzification of historical data (5) specifying the fuzzy logical relationship (6) grouping the fuzzy logical relationship (in step 5) (7) calculating its forecasting value.

2.2. *Fuzzy Time Series Markov Chain*

Markov Chain's Fuzzy Time Series forecasting procedure is as follows [9]:

Step 1. Collecting historical data (Yt).

Step 2. Defines the U universe set of data, with D_1 and D_2 being the corresponding positive numbers.

$$U = [D_{min} - D_1, D_{max} + D_2]$$

Step 3. Specify the number of fuzzy intervals.

Step 4. Defining the fuzzy set in the universe of discourse U, the Fuzzy A_i set declares the linguistic variable of the share price by $1 \leq i \leq n$.

Step 5. Fuzzification of historical data. If a time series data is included in the u_i interval, then that data is fuzzification into A_i.

Step 6. Specifies fuzzy logical relationship and Fuzzy Logical Relationships Group (FLRG).

Step 7. Calculate forecasting results

For time series data, using FLRG, a probability can be obtained from a state heading to the next state. So used markov probability transition matrix in calculating forecasting value, transition matrix size is $n \times n$. If state A_i transition to a state A_j and pass the state A_k, i, j= 1, 2, ..., n, then we will reap FLRG. The transition probability formula is as follows:

$$P_{ij} = \frac{M_{ij}}{M_i}, i, j = 1,2, \ldots, n, \quad (3)$$

with:
P_{ij}= probability of transition from state A_i to state A_j one step
M_{ij}= number of transitions from state A_i to state A_j one step
M_i = the quantity of data included in the A_i
The probability matrix R of all states can be written as follows:

$$R = \begin{bmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \cdots & P_{nn} \end{bmatrix} \quad (4)$$

Matrix R displays the transition of the entire system. If $F(t-1) = A_i$, then the procedure might be described withinside the A_i on the time of $(t-1)$, then the forecasting effects F (t) will be calculated using the $[P_{i1}, P_{i2}, \ldots, P_{in}]$ on the matrix R. Forecasting results F(t) is the weighted common value of the $m_1, m_2, \ldots, m_n$ (*midpoint* of $u_1, u_2, \ldots, u_n$). The forecasting output result value on F(t) can be determined the use of the subsequent rules:

a) Rule 1: if FLRG A_i is *one to one* (assume $A_i \rightarrow A_k$ wherein $P_{ik} = 1$ and $P_{ij} = 0, j \neq k$) then the forecasting value of F(t) is m_k the middle value of the u_k.

$$F(t) = m_k P_{ik} = m_k \quad (5)$$

b) Rule 2: if fuzzy logical relationship group (FLRG) A_i is *one to many* (assume $A_j \rightarrow A_1, A_2, \ldots, A_n$. $j = (1, 2, \ldots, n)$), while $Y(t-1)$ at the time of $(t-1)$ blanketed withinside the A_j then forecasting $F(t)$, is :

$$F(t) = m_1 P_{ji} + m_2 P_{i2} + \ldots + m_{j-1} P_{j(j-1)} + Y(t-1) P_{jj} + m_{j+1} P_{j(j+1)} + \ldots + m_n p_n \quad (6)$$

where: $m_1, m_2, \ldots, m_n$ is the middle value $u_1, u_2, \ldots, u_n$, Y(t-1) is the state value A_j at the time of t-1.

Step 8. Calculates the adjustment value (D_t) on the forecasting value. Here are the principles in calculating the value of adjustments:

a) If state A_i is related A_i, begining from state A_i on the time *t-1* declared as $F(t-1) = A_i$, and skilled a growing transition to the state A_j at the time t where (i < j) then the adjustment values are:

$$D_{t_1} = \left(\frac{l}{2}\right) \quad (7)$$

where l is the interval base.

b) If state A_i is related A_i, starting from state A_i at the time t-1 declared as F (t-1) = A_i, and decreasing transition to state A_j at the time t where (i > j) then the adjustment value is:

$$D_{t_1} = -\left(\frac{l}{2}\right) \quad (8)$$

c) If the transition starts from the A_i at the time t-1 declared as F (t-1) = A_i, and experienced a jump forward transition to the state A_{i+s} at the time t where ($1 \le s \le n-i$) then the adjustment value is:

$$D_{t_2} = \left(\frac{l}{2}\right)s, 1 < v < i \quad (9)$$

where s is the number of jump forward.

d) If the transition starts from the A_i at the time t - 1 declared as F (t – 1) = A_i, and experience a jump backward transition to the state A_{i-v} at the time t where ($1 \le v < i$) then the adjustment value is:

$$D_{t_2} = -\left(\frac{l}{2}\right)v, 1 \le v \le i \quad (10)$$

where v is the number of jump-backward.

Step 9. Calculate customized forecasting values

a) If FLRG A_i is *one to many* and state A_{i+1} accessible from the state A_i where A_i related A_i then the forecasting consequences grow to be

$$F'(t) = F(t) + D_{t_1} + D_{t_2} = F(t) + \left(\frac{l}{2}\right) + \left(\frac{l}{2}\right) \quad (11)$$

b) If FLRG A_i is *one to many* and state A_{i+1} accessible from the A_i where state A_i unrelated to A_i then the forecasting results become

$$F'(t) = F(t) + D_{t_2} = F(t) + \left(\frac{l}{2}\right) \quad (12)$$

c) If FLRG A_i is *one to many* and state A_{i-2} on hand from the state A_i where A_i do now no longer speak with A_i then the forecasting results become

$$F'(t) = F(t) - D_{t_2} = F(t) - \left(\frac{l}{2}\right) x\, 2 = F(t) - l \quad (13)$$

d) When v is *jump step*, the general form of forecasting results are:

$$F'(t) = F(t) \pm D_{t_1} \pm D_{t_2} = F(t) \pm \left(\frac{l}{2}\right) \pm \left(\frac{l}{2}\right) v \quad (14)$$

9

Markov Chain used Tsaur to foresee the exchange rate between the Taiwan and US dollars, which concluded that the markov chain had a totally small mean absolute percentage error value. Today many studies have used markov chain, including Lintang Afdianti Nurkhasanah, et al. In his research entitled comparison of fuzzy Chen method and fuzzy markov chain to are expacting inflation statistics in Indonesia which shows that based on the value of MSE obtained, fuzzy-Markov chain method was chosen as the best method because it produces the smallest mean square error [11] and Nurul Fitri, et al. in his research mentioned that batik sales forecasting from CV. Bintang Abadi uses 48 monthly sales data from January 2014 to December 2017 using fuzzy time series markov chain providing 1072 results and mape value = 22.4803% [12].

2.3. *Fuzzy time series Chen & Hsu*

Chen and Hsu strategies are carried out withinside the following steps [5]:

Step 1 :
- Determine the lag between n+1 and n data.
- Sum all the differences earned and then divided by the amount of data.
- Determine the length of the interval
- Determine the number of classes.
- Each class is symbolized by a universe set $U = U_1, U_2, U_3, \ldots, U_n$. according to the number of classes.
- Defines the set of universes of each class according to the interval length of each class.

Step 2 :
- Distribute all research data into each set of universes.
- Determine the amount of data included in each interval class.
- Redivided Interval.

Step 3 :
- Define the fuzzy set.
- Split fuzzy sets into sections.

Step 4 :
- Distribute fuzzy sets that have been formed into actual data tables.
- Forming Fuzzy Logical Relationship (FLR)

Step 5 :
- Determine the Difference among n-1 and n-2 data (Diff 1-2) and Difference between n-2 and n-3 data (Diff 2-3).
- Determine the difference between (Diff 1-2) – (Diff 2-3) which is then symbolized by DIFF.
- Specifies DIFF value x 2 + Data n-1 and DIFF/2 value + n-1 data.

In testing the accuracy of forecasting with the Fuzzy Time Series Chen and Hsu method, there are several rules that must be followed in determining the value of the forecast, namely:

Rule 1 :
- If the data analyzed does not have n-2 and n-3 data, then the middle value of fuzzy set A_j.
- If the data analyzed does not have n-3 data, then:
  a) if the difference between n-1 and n-2 > half the A_j interval, then the forecast value is expressed as upward 0.75 point interval A_j.
  b) if the difference between n-1 and n-2 = half the A_j interval then the forecast value is declared as middle value interval Aj.
  c) if the difference is n-1 and n-2 < half the interval of Aj then the forecast value is expressed as Downward interval Aj.

Rule 2 :
- If DIFF is postif value then:
  a) if the value (DIFF x 2 + Data n-1) is in the A_j interval then the forecast value is expressed as upward 0.75 point interval A_j.
  b) if the value (DIFF /2 + Data n-1) is in the A_j interval then the forecast value is expressed as downward 0.25 point interval A_j.

c) if point (a) and point (b) are not met then the forecast value is expressed with the Middle Value interval A_j.

Rule 3 :

- If DIFF is negative, then:

a) if the value (DIFF /2 + Data n-1) is in the A_j interval then the forecast value is expressed as downward 0.25 point interval A_j.

b) if the value (DIFF x 2 + Data n-1) is in the Aj interval then the forecast value is expressed as upward 0.75 point interval A_j.

c) If Point (a) and point (b) are not met, then the forecast value is expressed with the Middle Value of the A_j interval.

Chen & Hsu makes use of its method to forecast the weather, and the results have a high accuracy value. In its development, Chen & Hsu method has also been widely used in forecasting, among others to predict the value of Sharia Stock Exchange Index in Jakarta Islamic Index (JII) conducted by Rizka Zulfikar and Prihatin Ade Mayvita shows that Chen and Hsu techniques can be used to are expecting the trend of stock exchange indices that arise Jakarta Islamic Index with a reasonably correct of accuracy [13]. As well as being used in wulan anggraeni and Indra Suyahya's research entitled rupiah exchange rate prediction against the us dollar using the fuzzy time series chen &hsu method, which resulted in a low forecasting error rate [14].

3. Conclusion

The application of the Markov Chain and Chen & Hsu fuzzy time series is very simple and, based on several studies that have been described, proven that the method has a low error rate and has a high degree of accuracy. Because this study has not compared the calculations of Markov Chain and Chen &hsu in forecasting with the same data, it is necessary to forecast using the same data as the Markov Chain and Chen & Hsu methods so as to determine or choose the best method between the two.

References

[1] Montgomery C Douglas, Jennings L Cherly and M Kulahci 2015 *Introduction to Time Series Analysis and Forecasting Second Edition* (New York: Wiley-Interscience)
[2] Song Q and Chissom B S 1993 *Fuzzy sets Syst.* **54** 1 1
[3] Sullivan J and Woodall W 1994 *Fuzzy sets Syst.* **64** 279
[4] Chen S 1996 *Fuzzy sets Syst.* **81** 3 311
[5] Chen S and Hsu C 2004 *Inter. J. Apll. Scie. Eng.* **2** 3 234
[6] Yu H 2005 *J. Phys. A* **349** 609
[7] Cheng C H, Cheng G W and Wang J W 2008 *Exp. Syst. Appl.* **34** 2 1235
[8] Stevenson M, Porter J E 2009 *World Acad. Sci. Eng. Technol.* **55** 154
[9] Tsaur R 2012 *Int. J. Innov. Comput. Inf. Cont.* **8** 4931
[10] Jasim H T and Salim A G J 2012 *J. Appl. Appl. Math.* **7** 1 385
[11] Nurkhasanah, Suparti and Sudarno 2015 *J. Gauss.* **4** 4 917
[12] Fitriyah N, Faisol T Y 2018 *Proceed. Nat. Conf. Math. Sci. Educ. (NACOMSE)* 163
[13] Zulfikar R, Prihatini A M 2017 *Jurnal Penelitian Ilmu Ekonomi WIGA* **7** 108
[14] Suyahya I, Anggraeni W 2016 *Jurnal Sosio-E-Kons* **8** 1 24

Fuzzy time series Markov Chain and Fuzzy time series Chen & Hsu for forecasting

ORIGINALITY REPORT

17%	7%	15%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.researchgate.net Internet Source	2%
2	"Analysis and Design of Intelligent Systems using Soft Computing Techniques", Springer Science and Business Media LLC, 2007 Publication	2%
3	Yousif Alyousifi, Mahmod Othman, Rajalingam Sokkalingam, Ibrahima Faye, Petronio C. L. Silva. "Predicting Daily Air Pollution Index Based on Fuzzy Time Series Markov Chain Model", Symmetry, 2020 Publication	2%
4	Submitted to UIN Syarif Hidayatullah Jakarta Student Paper	1%
5	ejurnal.itats.ac.id Internet Source	1%
6	Shou-Hsiung Cheng, Shyi-Ming Chen, Wen-Shan Jian. "Fuzzy time series forecasting based on fuzzy logical relationships and	1%

similarity measures", Information Sciences, 2016
Publication

7	Shyi-Ming Chen. "Forecasting enrollments based on fuzzy time series", Fuzzy Sets and Systems, 1996 Publication	1 %
8	Frank van der Velde, Gwendid T. van der Voort van der Kleij, Marc de Kamps. "Lack of combinatorial productivity in language processing with simple recurrent networks", Connection Science, 2004 Publication	1 %
9	fmipa.unsrat.ac.id Internet Source	1 %
10	Submitted to Universitas Jember Student Paper	1 %
11	Hsien-Lun Wong. "An Evaluation of Comparison between Multivariate Fuzzy Time Series with Traditional Time Series Model for Forecasting Taiwan Export", 2009 WRI World Congress on Computer Science and Information Engineering, 03/2009 Publication	<1 %
12	Shyi-Ming Chen, Nai-Yi Wang. "Fuzzy Forecasting Based on Fuzzy-Trend Logical Relationship Groups", IEEE Transactions on	<1 %

Systems, Man, and Cybernetics, Part B (Cybernetics), 2010
Publication

13	Singh, S.R.. "A robust method of forecasting based on fuzzy time series", Applied Mathematics and Computation, 20070501 Publication	<1 %
14	Wang-Kun Chen. "Environmental Applications of Granular Computing and Intelligent Systems", Intelligent Systems Reference Library, 2011 Publication	<1 %
15	digilib.uinsby.ac.id Internet Source	<1 %
16	9lib.org Internet Source	<1 %
17	Advances in Intelligent Systems and Computing, 2016. Publication	<1 %
18	Melike Şah. "A New Trend Heuristic Time-Variant Fuzzy Time Series Method for Forecasting Enrollments", Lecture Notes in Computer Science, 2005 Publication	<1 %
19	Shyi-Ming Chen, G. M. T. Manalu, Jeng-Shyang Pan, Hsiang-Chuan Liu. "Fuzzy Forecasting Based on Two-Factors Second-Order Fuzzy-	<1 %

Trend Logical Relationship Groups and Particle Swarm Optimization Techniques", IEEE Transactions on Cybernetics, 2013
Publication

20	fs.unm.edu Internet Source	<1 %
21	Li, S.T.. "Deterministic fuzzy time series model for forecasting enrollments", Computers and Mathematics with Applications, 200706 Publication	<1 %
22	Sheng-Tun Li, Yi-Chung Cheng. "AN ENHANCED DETERMINISTIC FUZZY TIME SERIES FORECASTING MODEL", Cybernetics and Systems, 2009 Publication	<1 %
23	Charu Gupta, Amita Jain, Devendra Kumar Tayal, Oscar Castillo. "ClusFuDE: Forecasting low dimensional numerical data using an improved method based on automatic clustering, fuzzy relationships and differential evolution", Engineering Applications of Artificial Intelligence, 2018 Publication	<1 %

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		

Fuzzy time series Markov Chain and Fuzzy time series Chen & Hsu for forecasting

GRADEMARK REPORT

FINAL GRADE

/100

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6