Journal of Marine Science and Technology

Volume 19 | Issue 5 Article 6

LONG-TERM PREDICTIVE VALUE INTERVAL WITH THE FUZZY TIME SERIES

Ming-Tao Chou
Department of Aviation and Maritime Transportation Management, Chang Jung Christian University, Tainan County, Taiwan, R.O.C., mtchou@gmail.com

Follow this and additional works at: https://jmstt.ntou.edu.tw/journal

Part of the Engineering Commons

Recommended Citation
Chou, Ming-Tao (2011) "LONG-TERM PREDICTIVE VALUE INTERVAL WITH THE FUZZY TIME SERIES," *Journal of Marine Science and Technology*: Vol. 19: Iss. 5, Article 6.
DOI: 10.51400/2709-6998.2164
Available at: https://jmstt.ntou.edu.tw/journal/vol19/iss5/6

This Research Article is brought to you for free and open access by Journal of Marine Science and Technology. It has been accepted for inclusion in Journal of Marine Science and Technology by an authorized editor of Journal of Marine Science and Technology.

LONG-TERM PREDICTIVE VALUE INTERVAL WITH THE FUZZY TIME SERIES

Acknowledgements

The author gratefully acknowledges the helpful comments and suggestions of the reviewers, which have improved the presentation. This research work was partially supported by the National Science Council of the Republic of China under grant No. NSC 94-2211-E-309-002-.

This research article is available in Journal of Marine Science and Technology: https://jmstt.ntou.edu.tw/journal/vol19/iss5/6

LONG-TERM PREDICTIVE VALUE INTERVAL WITH THE FUZZY TIME SERIES

Ming-Tao Chou*

Key words: fuzzy time series, fuzzy forecasting, long-term predictive value interval.

ABSTRACT

This paper presents a new mathematical method to construct a fuzzy time series model of a system where the fuzzy support and interval values are used. This paper further presents an improved fuzzy time series model with a long-term predictive value interval and shows that the proposed definition can be used for long-term forecasting. The enrollment data of the University of Alabama (adopted by Song and Chissom, 1993) are used to demonstrate the proposed forecast model.

I. INTRODUCTION

Since its creation in 1965 by Zadeh [15], fuzzy set theory has enjoyed successful achievements both in theory advancement [15-17] and practical applications [1, 4, 6, 10]. In previous papers, the inadequacy of fuzzy time series models is found when fuzzy implications are used to express forecasting rules [1-17]. Several fuzzy time series models have been developed since Chen's [5] paper was published. The main purpose of this paper is to present a mathematical method to construct a fuzzy time series model for long-term prediction. This paper proposes a new method for long-term forecasting using the fuzzy time series based on Chou and Lee's method [5]. According to Chou and Lee's method [5], a long-term predictive value interval definition is added to improve attribution based on the definition of the fuzzy time series. This paper focuses on the enhancement of increasing and decreasing cases in Chou and Lee's framework. The current paper tackles the issues of improving the forecasting accuracy by controlling the uncertainties, and determining the support [7] of the fuzzy numbers. The university enrollment data used by Song and Chissom [2-5, 11-14] is used to demonstrate the proposed forecasting model. The remainder of this paper is organized as follows. Section 2 presents the fuzzy time series definition. In Section 3, the long-term predictive value interval using the fuzzy time series is presented. The forecasting of enrollment data is shown in Section 4. Finally, the conclusions are made in Section 5.

Paper submitted 11/28/08; revised 06/28/09; accepted 05/25/10. Author for correspondence: Ming-Tao Chou (e-mail: mtchou@gmail.com).
**Department of Aviation and Maritime Transportation Management, Chang Jung Christian University, Tainan County, Taiwan, R.O.C.*

II. FUZZY SETS AND FUZZY TIME SERIES

Song and Chissom [11-14] use discrete fuzzy sets to define the fuzzy time series. Chen [2] indicates Song and Chissom's method is too complicated, thus he proposes a simple method to compute the fuzzy relation. Lee and Chou [7] provide an objective interval method to define the universe of discourse. Liaw [9] investigates the nonstationary problem and employs the fuzzy trend technique to define the fuzzy time series. Various definitions and properties of fuzzy time series forecasting [2, 7, 9, 11, 12] are summarized as follows:

Definition 1 [8, 11] A fuzzy number on the real line $\Re$ is a fuzzy subset of $\Re$ that is normal and convex.

Definition 2 [11, 12] Let $Y(t)$ ($t = \ldots, 0, 1, 2, \ldots$), a subset of $\Re$, be the universe of discourse on which the fuzzy sets $f_i(t)$ ($t = 1, 2, \ldots$) are defined and $F(t)$ is the collection of $f_i(t)$ ($t = 1, 2, \ldots$). Then, $F(t)$ is called a fuzzy time series on $Y(t)$ ($t = \ldots, 0, 1, 2, \ldots$).

Definition 3 [11, 12] Let I and J be the index sets for $F(t-1)$ and $F(t)$ respectively. If for any $f_j(t) \in F(t)$ where $j \in J$, there exists $f_i(t-1) \in F(t-1)$ where $i \in I$ such that there exists a fuzzy relation $R_{ij}(t, t-1)$ and $f_j(t) = f_i(t-1) \circ R_{ij}(t, t-1)$ where '$\circ$' is the max-min composition, then $F(t)$ is said to be caused by only $F(t-1)$. Denote this as $f_i(t-1) \rightarrow f_j(t)$, or equivalently, $F(t-1) \rightarrow F(t)$.

Definition 4 [11, 12] If for any $f_j(t-1) \in F(\mathrm{t})$ where $j \in J$, there exists $f_i(t-1) \in F(t-1)$ where $i \in I$ and a fuzzy relation $R_{ij}(t, t = 1)$ such that $f_j(t) = f_i(t-1) \circ R_{ij}(t, t-1)$. Let $R(t, t-1) = \bigcup_{ij} R_{ij}(t, t-1)$ where $\bigcup$ is the union operator. Then, $R(t, t-1)$ is called the fuzzy relation between $F(t)$ and $F(t-1)$. Thus, we define this as the following fuzzy relational equation: $F(t) = F(t-1) \circ R(t, t-1)$.

Definition 5 [11, 12] Suppose that $R_1(t, t-1) = \bigcup_{ij} R^1_{ij}(t, t-1)$ and $R_2(t, t-1) = \bigcup_{ij} R^2_{ij}(t, t-1)$ are two fuzzy relations between

$F(t)$ and $F(t-1)$. If for any $f_j(t) \in F(t)$ where $j \in J$, there exist $f_i(t-1) \in F(t-1)$ where $i \in I$ and fuzzy relations $R^1_{ij}(t, t-1)$ and $R^2_{ij}(t, t-1)$ such that $f_i(t) = f_i(t-1) \circ R^1_{ij}(t, t-1)$ and $f_i(t) = f_i(t-1) \circ R^2_{ij}(t, t-1)$, then define $R_1(t, t-1) = R_2(t, t-1)$.

Definition 6 [11, 12] Suppose that $F(t)$ is only caused by $F(t-1)$ or $F(t-1)$ or $F(t-2)$ ... or $F(t-m)$ ($m > 0$). This relation can be expressed as the following fuzzy relational equation:

$$F(t) = F(t-1) \circ R_0(t, t-m)$$

This equation is called the first-order model of $F(t)$.

Definition 7 [11, 12] Suppose that $F(t)$ is simultaneously caused by $F(t-1)$, $F(t-2)$, …, and $F(t-m)$ ($m > 0$). This relation can be expressed as the following fuzzy relational equation:

$$F(t) = (F(t-1) \times F(t-2) \times \ldots \times F(t-m)) \circ R_a(t, t-m)$$

This equation is called the mth order model of $F(t)$.

Definition 8 [2] $F(t)$ is a fuzzy time series if $F(t)$ is a fuzzy set. The transition is denoted as $F(t-1) \to F(t)$.

Definition 9 [7] The universe of discourse $U = [D_L, D_U]$ is defined such that $D_L = D_{\min} - st_\alpha(n)/\sqrt{n}$ and $D_U = D_{\max} + st_\alpha(n)/\sqrt{n}$ when $n \le 30$ or $D_L = D_{\min} - \sigma Z_\alpha/\sqrt{n}$ and $D_U = D_{\max} - \sigma Z_\alpha/\sqrt{n}$ when $n > 30$, where $t_\alpha(n)$ is the $100(1-\alpha)$ percentile of the t distribution with n degrees of freedom and z_α is the $100(1-\alpha)$ percentile of the standard normal distribution, that is, if Z is a $N(0, 1)$ distribution, then $P(Z \ge z_\alpha) = \alpha$.

Definition 10 [7] Assuming that there are m linguistic values under consideration, let A_i be the fuzzy number that represents the i[th] linguistic value of the linguistic variable where $1 \le i \le m$. The support of A_i is defined to be

$$\begin{cases} D_L + (i-1)\dfrac{D_U - D_L}{m},\ D_L + \dfrac{i(D_U - D_L)}{m},\ 1 \le i \le m-1 \\ D_L + (i-1)\dfrac{D_U - D_L}{m},\ D_L + \dfrac{i(D_U - D_L)}{m},\ i = m. \end{cases}$$

Definition 11 [9] For a test H_0: *nonfuzzy trend* against H_1: *fuzzy trend*, where the critical region $C^* = \{C \mid C_2^k + C_2^{n-k} > C_\lambda = C_2^n \times (1-\lambda)\}$ and the initial value of the significant level α is 0.2.

III. THE STATIC LONG-TERM PREDICTIVE VALUE INTERVAL

In this section, a useful method is proposed to forecast the long-term predictive value interval using a fuzzy time series. Generally speaking, this method is an extended model of the method proposed by Chou and Lee [5, 7].

Definition 12 Let $d(t)$ be a set of real numbers: $d(t) \subseteq R$. An upper interval for $d(t)$ is a number b such that $x \le b$ for all $x \in d(t)$. $d(t)$ is said to be an interval above if $d(t)$ has an upper interval. A number, max, is the maximum of $d(t)$ if max is an upper interval for $d(t)$ and max $\in d(t)$.

Definition 13 Let $d(t) \subseteq R$. The least upper interval of $d(t)$ is a number $\overrightarrow{\max}$ satisfying:

(1): $\overrightarrow{\max}$ is an upper interval for $d(t)$: $x \le \overrightarrow{\max}$ for all $x \in d(t)$.

(2): $\overrightarrow{\max}$ is the least upper interval for $d(t)$, that is $x \le b$ for all $x \in d(t) \Rightarrow \overrightarrow{\max} \le b$.

Definition 14 Let $d(t)$ be a set of real numbers: $d(t) \subseteq R$. A lower interval for $d(t)$ is a number b such that $x \ge b$ for all $x \in d(t)$. $d(t)$ is said to be an interval below if $d(t)$ has a lower interval. A number, min, is the minimum of $d(t)$ if min is a lower interval for $d(t)$ and min $\in d(t)$.

Definition 15 Let $d(t) \subseteq R$. The least lower interval of $d(t)$ is a number $\overleftarrow{\min}$ satisfying:

(1): $\overleftarrow{\min}$ is a lower interval for $d(t)$: $x \ge \overleftarrow{\min}$ for all $x \in d(t)$.

(2): $\overleftarrow{\min}$ is the least lower interval for $d(t)$, that is $x \ge b$ for all $x \in d(t) \Rightarrow \overleftarrow{\min} \le b$.

Definition 16 The long-term predictive value interval, $(\overleftarrow{\min}, \overrightarrow{\max})$ is called the static long-term predictive value interval.

Following Definitions 12-16, $(\overleftarrow{\min}, \overrightarrow{\max})$ can be obtained. In summary, the long-term predictive value interval for $d(t)$ is given by $(\overleftarrow{\min}, \overrightarrow{\max})$. The stepwise procedure of the proposed method consists of the following steps and a flow diagram is shown in Fig. 1.

Step 1. Let $d(t)$ be the data under consideration and $F(t)$ be the fuzzy time series. Following Definition 11, a difference test is used to understand whether or not the information is in a stable state. Recursion is performed until the information is in a stable state, where the critical region is $C^* = \{C \mid C_2^k + C_2^{n-k} > C_\lambda = C_2^n \times (1-\lambda)\}$.

Step 2. Determine the universe of discourse $U = [D_L, D_U]$.

Step 3. Define A_i by letting its membership function be as follows:

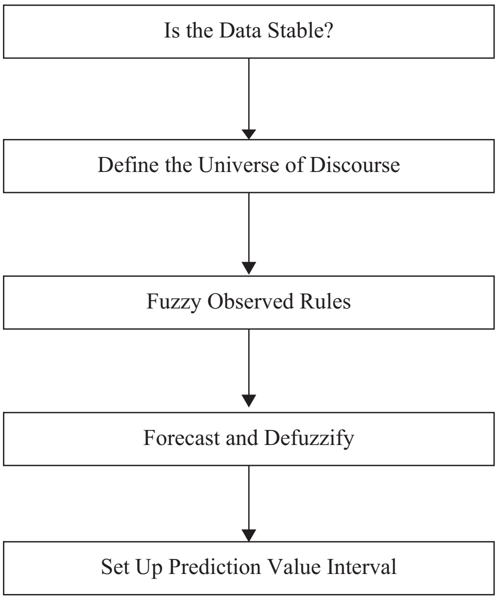

Fig. 1. The procedure of the proposed model.

$$u_{A_i}(x) = \begin{cases} 1 \ \text{ for } x \in [D_L + (i-1)\dfrac{D_U - D_L}{m}, D_L + \dfrac{i(D_U - D_L)}{m}) \\ \text{where } 1 \le i \le m-1; \\ 1 \ \text{ for } x \in [D_L + (i-1)\dfrac{D_U - D_L}{m}, D_L + \dfrac{i(D_U - D_L)}{m}] \\ \text{where } i = m; \\ 0, \ \text{otherwise.} \end{cases}$$

Step 4. Then, $F(t) = A_i$ if $d(t) \in \text{supp}(A_i)$, where supp $(\cdot)$ denotes the support.

Step 5. Derive the transition rule from period $t-1$ to t and denote it as $F(t-1) \rightarrow F(t)$. Aggregate all transition rules. Let the set of rules be $R = \{r_i \mid r_i : P_i \rightarrow Q_i\}$.

Step 6. The value of $d(t)$ can be predicted using the fuzzy time series $F(t)$ as follows:
Let $T(t) = \{r_j \mid d(t) \in \text{supp}(P_j), \textit{where } r_j \in R\}$ be the set of rules fired by $d(t)$, where $\text{supp}(P_j)$ is the support of P_j. Let $\overline{\text{supp}(P_j)}$ be the median of $\text{supp}(P_j)$. The predicted value of $d(t)$ is $\sum_{r_j \in T(t-1)} \frac{\overline{\text{supp}(Q_j)}}{|T(t-1)|}$.

Step 7. The long-term predictive value interval for $d(t)$ is given as $(\overleftarrow{\min}, \overrightarrow{\max})$.

IV. ENROLLMENT FORECASTING

Enrollment forecasting will be used to demonstrate how the proposed fuzzy time series model can be applied to forecast long-term data. The walk through of enrollment forecasting is presented as follows.

Table 1. Fuzzy historical enrollment data and forecasted enrollment.

Year	Actual	Fuzzified enrollments	The forecasted results
1971	13,055	A1	
1972	13,563	A2	14,025
1973	13,867	A2	14,568
1974	14,696	A3	14,568
1975	15,460	A3	15,654
1976	15,311	A3	15,654
1977	15,603	A3	15,654
1978	15,861	A4	15,654
1979	16,807	A5	16,197
1980	16,919	A5	17,283
1981	16,388	A4	17,283
1982	15,433	A3	16,197
1983	15,497	A3	15,654
1984	15,145	A3	15,654
1985	15,163	A3	15,654
1986	15,984	A4	15,654
1987	16,859	A5	16,197
1988	18,150	A6	17,283
1989	18,970	A7	18,369
1990	19,328	A7	19,454
1991	19,337	A7	19,454

Source: [7].

1. Forecasting of Enrollment:

Step 1. Let $d(t)$ be the historical data under consideration and $F(t)$ be the fuzzy time series. Following Definition 11, a difference test is used to understand whether or not the information is in a stable state. Recursion is performed until the information is in a stable. Since = $C'' = \{C \mid C = C_2^4 + C_2^{21-4}\} = 142 < \{C \mid C_2^{21} \times (1-0.2)\} = 168$, that the information is in a stable state is not rejected.

Step 2. From Definition 9, the discourse $U = \{D_L, D_U\}$. From Table 1, $D_{\min} = 13{,}055$, $D_{\max} = 19{,}337$, $s = 1{,}757$, and $n = 21$ can be obtained. Let $\alpha = 0.05$. Since n is less than 30, the Student-t distribution with 21 degrees of freedom is used as a substitute for the normal distribution. Thus, $t_\alpha(n) = t_{0.05}(21) = 1.721$, $D_L = D_{\min} - st_\alpha/\sqrt{n} \approx 12{,}396$, and $D_U = D_{\max} + st_\alpha/\sqrt{n} \approx 19{,}996$. That is, $U = [12{,}396, 19{,}996]$.

Step 3. Assume that the following linguistic values are under consideration: extremely few, very few, few, some, many, very many, and extremely many. According to Definition 10, the supports of fuzzy numbers that represent the linguistic values are given by

Table 2. Fuzzy transitions derived from Table 1.

$r_1 : A_1 \to A_2$	$r_6 : A_4 \to A_3$	$r_{11} : A_6 \to A_7$
$r_2 : A_2 \to A_2$	$r_7 : A_4 \to A_5$	$r_{12} : A_7 \to A_7$
$r_3 : A_2 \to A_3$	$r_8 : A_5 \to A_4$	
$r_4 : A_3 \to A_3$	$r_9 : A_5 \to A_6$	
$r_5 : A_3 \to A_4$	$r_{10} : A_5 \to A_5$	

Source: [7].

$$u_{A_i}(x) = \begin{cases} 1 \text{ for } x \in [12{,}396 + (i-1)(1{,}086), 12{,}396 + i(1{,}086)) \\ \text{where } 1 \le i \le m-1; \\ 1 \text{ for } x \in [12{,}396 + (i-1)(1{,}086), 12{,}396 + i(1{,}086)] \\ \text{where } i = m; \\ 0 \text{ otherwise.} \end{cases}$$

where A_1 = extremely few, A_2 = very few, A_3 = few, A_4 = some, A_5 = many, A_6 = very many, and A_7 = extremely many. Thus, the supports are supp(A_1) = [12,396, 13482], supp(A_2) = [13,482, 14,568], supp(A_3) = [14,568, 15,654], supp(A_4) = [15,654, 16,740], supp(A_5) = [16,740, 17,826], supp(A_6) = [17,826, 18,912], and supp(A_7) = [18,912, 19,996].

Step 4. The fuzzy time series $F(t)$ is given by $F(t) = A_i$ when $d(t) \in$ supp(A_i). Therefore,

$F(1971) = A_1$, $F(1972) = A_2$, $F(1974) = A_3$, $F(1979) = A_5$, …, and $F(1989) = A_7$

A comparison between actual enrollment data and the fuzzy enrollment data is shown in Table 1.

Step 5. The transition rules are derived from Table 1. For example, $F(1977) \to F(1978)$ is $A_3 \to A_4$. All transition rules obtained from Table 1 are shown in Table 2.

Step 6. The forecasting results from 1972 to 1991 are shown in Table 1.

Step 7. The long-term predictive value interval for $d(t)$ is given as (14,025, 19,454).

2. Discussion

One of the major limitations of the existing fuzzy time series forecasting models [2-5, 7, 11-14] is that they can only provide a single-point forecast value. The current paper proposes a method which is a composite of the fuzzy support [7] and the stability concept of fuzzy time series [9]. This method not only provides a more objective interval setup technique, [7] but also increases the forecasting applicability [2-5, 7, 11-14].

It is demonstrated that the proposed time series model provides an adequate fit to the data and values for the derived fuzzy transitions can be forecasted by using Definitions 12-16. To forecast the long-term predictive value interval for $d(t)$, the min-max interval, is given by $(\overleftarrow{\min}, \overrightarrow{\max}) = (14{,}025, 19{,}454)$. According to the $(\overleftarrow{\min}, \overrightarrow{\max})$, it is clear that the proposed method can provide better forecasting information than other methods [2-5, 7, 11-14].

V. CONCLUSIONS

In this paper, a long-term predictive value interval model has been developed for the fuzzy time series. This model helps to minimize the uncertainties of the fuzzy numbers and handle fuzzy-trend/non-fuzzy-trend [9]. The method is examined by forecasting the enrollment of a university from its enrollment data from which the min-max interval, $(\overleftarrow{\min}, \overrightarrow{\max})$, is obtained. In summary, the proposed method is as accurate as the methods used in the past [2, 3, 7, 11-14], and is more robust and systematic in forecasting.

ACKNOWLEDGMENTS

The author gratefully acknowledges the helpful comments and suggestions of the reviewers, which have improved the presentation. This research work was partially supported by the National Science Council of the Republic of China under grant No. NSC 94-2211-E-309-002-.

REFERENCES

1. Chen, S.-H., Wang, C.-C., and Chang, S.-M., "Fuzzy economic production quantity model for items with imperfect quality," *International Journal of Innovative Computing, Information and Control*, Vol. 3, No. 1, pp. 85-95 (2007).
2. Chen, S.-M., "Forecasting enrollments based on fuzzy time series," *Fuzzy Sets and Systems*, Vol. 81, No. 3, pp. 311-319 (1996).
3. Chen, S.-M., "Forecasting enrollments based on high order fuzzy time series," *Cybernetics and Systems: An International Journal*, Vol. 33, No. 1, pp. 1-16 (2002).
4. Chou, M.-T., "A fuzzy time series model to forecast the BDI," *IEEE Proceeding of the Fourth International Conference on Networked Computing and Advanced Information Management*, Gyeongju, Korea, pp. 50-53 (2008).
5. Chou, M.-T. and Lee, H.-S., "Increasing and decreasing with fuzzy time series," *Joint Conference on Information Sciences*, Kaohsiung, Taiwan, pp. 1240-1243 (2006).
6. Han, T.-C., Chung, C.-C., and Liang, G.-S., "Application of fuzzy critical path method to operation systems," *Journal of Marine Science and Technology*, Vol. 14, No. 3, pp. 139-146 (2006).
7. Lee, H.-S. and Chou, M.-T., "Fuzzy forecast based on fuzzy time series," *International Journal of Computer Mathematic*, Vol. 81, No. 7, pp. 781-789 (2004).
8. Liang, M.-T., Wu, J.-H., and Liang, G.-S., "Applying fuzzy mathematics to evaluating the membership of existing reinforced concrete bridges in Taipei," *Journal of Marine Science and Technology*, Vol. 8, No. 1, pp. 16-29 (2006).
9. Liaw, M.-C., *The Order Identification of Fuzzy Time Series, Models Construction and Forecasting*, Ph.D. Dissertation, Department of Statistics, National Chengchi University, Taiwan (1997).
10. Lin, S.-C., Liang, G.-S., and Lee, K.-L., "Applying fuzzy analytic hierarchy process in location mode of international logistics on airports competition evaluation," *Journal of Marine Science and Technology*, Vol. 14, No. 1, pp. 25-38 (2006).
11. Song, Q. and Chissom, B. S., "Fuzzy time series and its models," *Fuzzy Sets and Systems*, Vol. 54, No. 3, pp. 269-277 (1993).

12. Song, Q. and Chissom, B. S., “Forecasting enrollment with fuzzy time series-Part I,” *Fuzzy Sets and Systems*, Vol. 54, No. 1, pp. 1-9 (1993).
13. Song, Q. and Chissom, B. S., “Forecasting enrollment with fuzzy time series-Part II,” *Fuzzy Sets and Systems*, Vol. 62, No. 1, pp. 1-8 (1994).
14. Song, Q., Leland, R. P., and Chissom, B. S., “Fuzzy stochastic time series and its models,” *Fuzzy Sets and Systems*, Vol. 88, No. 3, pp. 333-341 (1997).
15. Zadeh, L. A., “Fuzzy set,” *Information and Control*, Vol. 8, No. 3, pp. 338-353 (1965).
16. Zadeh, L. A., “Similarity relations and fuzzy orderings,” *Information Sciences*, Vol. 3, No. 2, pp. 177-200 (1971).
17. Zadeh, L. A., “Outline of a new approach to the analysis of complex system and decision process,” *IEEE Transactions on System, Man and Cybernetics*, Vol. 3, No. 1, pp. 28-44 (1973).