2009 International Conference on Machine Learning and Computing
IPCSIT vol.3 (2011) © (2011) IACSIT Press, Singapore

Information Fusion Technique for Fuzzy Time Series Model

Jia-Wen Wang [1+], Jing-Wei Liu [2]

[1]Department of Electronic Commerce Management, Nanhua University, Taiwan

[2] Department of Information Management, National Yunlin University of Science & Technology, Taiwan

Abstract. This paper proposes a high order multiple attribute fuzzy time series model, which corporate a clustering method and order weighted averaging (OWA) operator. The proposed model can deal with (1) lacking persuasiveness in determining the universe of discourse and the linguistic length of intervals, and (2) only one attribute is usually considered in forecasting not multiple attributes. Furthermore, we can according to different situation α to adjust high order forecasting value. To verify the proposed model, we use the yearly data on enrollments at the University of Alabama and TAIFEX (Taiwan Futures Exchange) as experimental datasets. Finally, this paper compares forecasting performances of proposed methods with Hang et al.'s [1] and Cheng et al.'s [2] models, the results of empirical analysis conclude that the proposed model surpasses in accuracy the listing models.

Keywords: Fuzzy time series, Order Weighted Averaging (OWA), Fuzzy clustering, Fuzzy logic relationship (FLR)

1. Introduction

In recent years, time-series models have utilized the fuzzy theory to solve various forecasting problems, such as university enrolment forecasting [3], financial forecasting [1, 4] and temperature forecasting [5]. Many researchers have proposed different methods to predict stock price based on fuzzy time series. In 1993, Song and Chissom[6] proposed the first fuzzy time series method, where the definitions, the time-invariant model[4] and the time-variant model[1] of fuzzy time series were presented. The following research proposed simple calculation methods to get a higher forecasting accuracy [3]. Huang and Yu utilized two methods, distribution-based and average-based length, to set the linguistic intervals of the universe of discourse to improve forecasting accuracy [7]. From the literature above, there are drawbacks: (1) Many factors in stock markets, such as financial report, macro economical data, and the fluctuations of foreign stock markets, influence practically the decisions of stock investors, and, therefore, multiple attributes should be considered in forecasting model. (2) In stock market forecasting, it is not reasonable to partition the universe of the discourse for dataset because the distribution of stock price is not uniform.

Therefore, this paper proposes a multiple attribute fuzzy time series model, which corporate a clustering method and order weighted averaging (OWA) operator. The proposed model can deal with (1) lacking persuasiveness in determining the universe of discourse and the linguistic length of intervals, and (2) only one attribute is usually considered in forecasting not multiple attributes. Furthermore, we can according to different situation α to adjust forecasting value.

The rest of this paper is organized as follows: section 2 introduces the related literature of fuzzy time-series model, Order Weighted Averaging (OWA) operator; in section 3, demonstrates the proposed model and algorithm; section 4 evaluates the performance of the proposed model; and, finally, conclusions are presented in section 5.

+ Corresponding author. Tel.: + 052721001#56237; fax: +886-5-2427197.
E-mail address: jwwang@mail.nhu.edu.tw.(Jia-Wen Wang)

2. Related work

In this section, several related literatures including Fuzzy time-series, order weighted averaging (OWA), and Evaluation criteria are briefly reviewed.

2.1. Fuzzy Time-series

Song and Chissom [6] proposed a Fuzzy Time Series model to deal with the problems involving human linguistic terms [8-10]. In the following research, they continued to discuss the difference between time-invariant and time-variant models[6, 11]. Besides the above researchers, Chen presented a method to forecast the enrollments of the University of Alabama based on fuzzy time series [3].

Over the past fourteen years, many fuzzy time-series models have been proposed by following Song and Chissom's definitions [6, 11, 12]. Among these models, Chen's model is very conventional one because of easy calculations and good forecasting performance [3, 5]. Therefore, Song and Chissom's definitions and Chen's algorithm are used for illustrations as follows:

Definition 1: fuzzy time-series

Let $Y(t)(t=\ldots,0,1,2,\ldots)$, a subset of real numbers, be the universe of discourse by which fuzzy sets $f_j(t)$ are defined. If $F(t)$ is a collection of $f_1(t), f_2(t)\cdots$ then $F(t)$ is called a fuzzy time-series defined on $y(t)$.

Definition 2: fuzzy time-series relationships

Assuming that $F(t)$ is caused only by $F(t-1)$, then the relationship can be expressed as: $F(t)=F(t-1)*R(t,t-1)$, which is the fuzzy relationship between $F(t)$ and $F(t-1)$, where * represents as an operator. To sum up, let $F(t-1)=A_i$ and $F(t)=A_j$. The fuzzy logical relationship between $F(t)$ and $F(t-1)$ can be denoted as $A_i \rightarrow A_j$ where A_i refers to the left-hand side and A_j refers to the right-hand side of the fuzzy logical relationship. Furthermore, these fuzzy logical relationships can be grouped to establish different fuzzy relationships.

2.2. Information Fusion Techniques: Order Weighted Averaging (OWA)

Yager proposed an order weighted averaging (OWA) operator which had the ability to get optimal weights of the attributes based on the rank of these weighting vectors after aggregation process [13]. An OWA operator of dimension n is a mapping f : $R^n \rightarrow R$, that has an associated weighting vector W=$[w_1, w_2, \ldots\ldots, w_n]^T$ with the following properties: $W_i \in [0,1]$ for $i \in I = \{1,2,\ldots,n\}$ and $\sum_{i \in I} W_i = 1$, Such that

$$f(a_1, a_2, \ldots, a_n) = \sum_{i \in I} W_i b_i \tag{1}$$

where b_i is the ith largest element in the collection. Thus, it satisfies the relation $Min_i[a_i] \le f(a_1, a_2, \ldots, a_n) \le Max_i[a_i]$.

Fuller and Majlender [14] transform Yager's OWA equation to a polynomial equation by using Lagrange multipliers. According to their approach, the associated weighting vector can be obtained by (2) ~ (4)

$$\ln w_j = \frac{j-1}{n-1}\ln w_n + \frac{n-j}{n-1}\ln w_1 \Rightarrow w_j = \sqrt[n-1]{w_1^{n-j} w_n^{j-1}} \tag{2}$$

and

$$w_1[(n-1)\alpha + 1 - n w_1]^n = [(n-1)\alpha]^{n-1}\,[((n-1)\alpha - n) w_1 + 1] \tag{3}$$

if $w_1 = w_2 = \ldots\ldots = w_n = \frac{1}{n} \Rightarrow$ disp(W) = $\ln n$ ($\alpha = 0.5$) then

$$w_n = \frac{((n-1)\alpha - n) w_1 + 1}{(n-1)\alpha + 1 - n w_1} \tag{4}$$

where W_i is weight vector, N is number of attributes, and α is the situation parameter.

2.3. Evaluation criteria

MSE (Mean Squared Error) is used to measure performance. MSE is defined as eq(5).

$$MSE = \frac{\sum_{i=1}^{n}(Actual(t) - Forecast(t))^2}{n} \tag{5}$$

where actual(t) is the actual observations on time t; forecast(t) is the forecasting value; and n is the number of periods for forecasts.

3. The Proposed model

In order to deal with multiple attribute prediction and improve forecasting accuracy, hence, this paper proposes a hybrid model: (1) clustering method (fuzzy c-mean [15]) for partitioning process, and (2) OWA operator for high order adjusting.

3.1. OWA weight

In this section, we utilized an algorithm in order to calculate the OWA weights [16]. The main idea is from equation (2) ~ (4)[16]. The main procedure of the program is simply presented as following (as fig. 1):

For different α and *n*, we can get different OWA weight. /* n is the number of attributes; α is the situation parameter */

OWA (n, α)

If α < 0.5

Then α = 1- α

If α > 0.5

Then $w_1[(n-1)\alpha + 1 - nw_1]^n = [(n-1)\alpha]^{n-1}\,[((n-1)\alpha - n)w_1 + 1]$ //Calculate W_1

$w_n = [((n-1)\alpha - n)\,w_1 + 1]/[(n-1)\alpha + 1 - n\,w_1]$ //Calculate W_n

For *i* = 2 to *n*-1 do

$w_i = \sqrt[n-1]{w_1^{n-i}\,w_n^{i-1}}$ //Calculate W_i

Figure 1. Algorithm of OWA [16]

For example, we compute OWA weight under n=3 and situation α=0.5~1.0 in table1

Table 1. The w_1 *~ w_3* values for different situation values α

	α=0.5	α=0.6	α=0.7	α=0.8	α=0.9	α=1
w_1^*	0.333333	0.438355	0.553955	0.681854	0.826294	1
w_2^*	0.333333	0.323242	0.291992	0.23584	0.146973	0
w_3^*	0.333333	0.238392	0.153999	0.081892	0.026306	0

3.2. Algorithm of the proposed method

In this section, we utilize the yearly data on enrollments at the University of Alabama to present the proposed algorithm.

Step 1: Select attributes

According to research problem, select attributes by domain experts. From table 2, the yearly enrollments of the University of Alabama only have one attribute; hence, we use enrollments as main attribute.

Step 2: Computing difference of each period

In this step, we calculate difference of each period. From table 1, difference of 1971 and 1972 is 508(=13563-13055). Therefore, we can get difference of each period in table 2.

Step 3: Cluster multi-attributes time series S(t)

Supposed a time series S(t) with n observations of m attributes, an appropriate fuzzy clustering procedure is selected to cluster time series S(t) into c ($2 \le c \le n$) clusters in this step. In this study, FCM is chosen because it is one of the fuzzy clustering methods.

According to (Miller, 1956), seven (linguistic value) is utilized as the number of clusters for demonstration to correspond with the limitation of human cognition in shorten memory. Hence, we use difference of each period into FCM by equation (3).

Step 4: Rank each cluster and fuzzify the time series S(t) as fuzzy time series F(t)

In table 2, we ranked each of cluster center value and got linguistic value.

Step 5: Fuzzify the historical data

Firstly, define the fuzzy set, L1 L2… L7,, on the universe of discourse.

Secondly, find out the degree of each observation value belonging to each Li (i=1,…,7). If the maximum membership of the observation value is under L7, then the fuzzified observation value is labeled as L7. Lastly, convert each stock price in training dataset to corresponding linguistic values, L7. From table 2, the maximum membership of the difference of 1971 and 1972 is 0.999773 in fifth cluster.

Step 6: Establish fuzzy logic relationships and the fuzzy logic relationships groups

Construct fuzzy logical relationship between two consecutive linguistic values such as $A_i \to A_j$, where A_i is called the LHS (left-hand side) and A_j the RHS (right-hand side) of the FLR. Then, we establish fuzzy logical relationship groups. For instance, Ai include fuzzy logical relationship $A_i \to A_j$, $A_i \to A_k$, and $A_i \to A_l$, hence, the fuzzy logical relationship group is $A_i \to A_j, A_k, A_l$.

Step 7: Defuzzify and compute forecast value

Suppose F(t) = Li, the forecasting of F(t+1) is conducted using the following rules.

Rule 1: According to the Naïve Forecasting Principle, if the fuzzy relationship group of Li is empty, such as Li→empty, the forecasting of F(t+1) is Li.

Rule 2: If the fuzzy relationship group of Li is one-to-one, such as Li→Lx, then the forecasting of F(t+1) is Lx.

Rule 3: If the fuzzy relationship group of Li is one-to-many, such as Li→L2,…,Lk, then the forecasting of F(t+1) is supposed that each linguistic values is of the equal weight. Therefore, arithmetic mean is used as the forecasted value. The equation is as following:

$$\frac{\mathrm{Def}(L_1)+\mathrm{Def}(L2)+\ldots+\mathrm{Def}(L_k)}{k} \tag{6}$$

where Def( ) is denoted as the defuzzified value of that linguistic values and k is number of linguistic values.

From table 2, if we want to forecast the enrollments in 1973, fuzzy relation group L4→L3 , L4,, L6 is inferred. The forecasted value = (center of L3 + center of L4+ center of L6)/3=(297+24+840)/3=387

Table 2 Results of forecasting difference(part)

Year	Enrollments	difference	Linguistic value	Forecasted	Year	Enrollments	difference	Linguistic value	Forecasted
1971	13055				…	…	…	…	…
1972	13563	508	5	297	1990	19328	358	4	387
1973	13867	304	4	387	1991	19337	9	3	230
1974	14696	829	6	612	1992	18876	-461	2	-465

Step 8: Calculate adjusted forecasting values based on OWA operator

From 3.1, the attributes ordering and number are known. For computing the aggregated values, we multiply the values of attributes ordering by corresponding to the weights of OWA. For example, we calculate forecasting value in 1975 under α=0.5 and n=3. Forecasting_value1975=15460+612*0.333333+387*0.333333+297*0.333333= 15892

4. Verifications and comparisons

In this section, to illustrate how the proposed model performs fuzzy time series forecasting, practically collected TAIEX (Taiwan Stock Exchange Capitalization Weighted Stock Index) futures are employed as experiment datasets. Finally, we use MSE as performance indicator to compare proposed model with Hang et al.'s [1], and Cheng et al.'s [2] models.

Forecasting TAIEX Futures

In this case, we practically collect TAIEX futures (2004/1/1~2004/12/15) as experiment datasets. From this datasets have 208 records, which are utilized 248 records (2004/1/1~2004/11/19) as training datasets, others are testing datasets (18 records, 2004/11/22~2004/12/15). To examine the improvement in performance, three fuzzy time-series models, Hang et al.'s [1], and Cheng et al.'s [2]] models, are employed as comparison models. The comparison result is shown in Table 3, MSE of the proposed model are 496.87. Clearly, the performance of the proposed model surpasses the listing models.

Table 3 Results and comparisons for TAIFEX forecasting

		Type 2 method [1]	Cheng et al.'s[2]	The proposed model
MSE	Training data	21950.16	20431.44	488.81
	Testing data	32252.22	8297.627	496.87

5. Conclusions

In this paper, a high order multiple attribute fuzzy time series method is proposed which employs fuzzy c-mean clustering method and OWA operator. The proposed model can overcome (1) lacking persuasiveness in determining the universe of discourse and the linguistic length of intervals, and (2) only one attribute is usually considered in forecasting not multiple attributes. Besides, we utilize different situation α to adjust high order forecasting values. Finally, from the compare result, we can see that the proposed model outperforms the listing models. Based on the verification results, we conclude that the research goal has been reached.

6. Acknowledgements

The first author gratefully appreciates the financial support from National Science Council, Taiwan, ROC under contract NSC 97-2221-E-343-004

7. References

[1] K. Huarng and H.-K. Yu, "A type 2 fuzzy time series model for stock index forecasting," *Physica A*, vol. 353, pp. 445-462, 2005.

[2] Cheng, C.-H., Cheng, G.-W., & Wang, J.-W. (2008). Multi-attribute fuzzy time series method based on fuzzy clustering. *Expert Systems with Applications*, 34(2), 1235-1242.

[3] S. M. Chen, "Forecasting enrollments based on fuzzy time series," *Fuzzy Sets Syst,* vol. 81, pp. 311-319, 1996.

[4] K. Huarng, "Heuristic models of fuzzy time series for forecasting," *Fuzzy Sets Syst*, vol. 123, pp. 369–386, 2001.

[5] S. M. Chen and J. R. Hwang, "Temperature prediction using fuzzy time series," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 30, pp. 263–275, 2000.

[6] Q. Song and B. S. Chissom, "Fuzzy time series and its models," *Fuzzy Sets Syst*, vol. 54, pp. 269-277, 1993.

[7] K. Huarng and H.-K. Yu, "An N-th order heuristic fuzzy time series model for TAIEX forecasting.," *International Journal of Fuzzy Systems,* vol. 5, pp. 247-253, 2008.

[8] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning, Part I," *Inform. Sci,* vol. 8, pp. 199–249, 1975.

[9] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning II," *Information Science*, vol. 8, pp. 301-357, 1975.

[10] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning III," *Information Science,* vol. 19, pp. 43-80, 1976.

[11] Q. Song and B. S. Chissom, "Forecasting enrollments with fuzzy time-series - Part II," *Fuzzy Sets and Systems,* vol. 62, pp. 1-8, 1994.

[12] Q. Song and B. S. Chissom, "Forecasting enrollments with fuzzy time-series - Part I," *Fuzzy Sets and Systems*, vol. 54, pp. 1-9, 1993.

[13] R. R. Yager, "On ordered weighted averaging aggregation operators in mult-icriteria decision making," *IEEE Trans. Systems Man and Cybernetics*, pp. 183-190, 1988.

[14] R. Fuller and P. Majlender, "An analytic approach for obtaining maximal entropy OWA operator weights," *Fuzzy Sets and Systems*, vol. 124, pp. 53-57, 2001.

[15] J. C. Bezdek, Pattern recognition with fuzzy objective function algorithms. *NY: Plenum Press,* 1981.

[16] Cheng, C.-H., Cheng, Jia-Wen Wang, Ming-Chang Wu. (2009). OWA-weighted based clustering method for classification problem. *Expert Systems with Applications,* 36(3), 4988-4995.