International Journal of Uncertainty,
Fuzziness and Knowledge-Based Systems
Vol. 11, No. 6 (2003) 671–690

A NEW APPROACH OF BIVARIATE FUZZY TIME SERIES ANALYSIS TO THE FORECASTING OF A STOCK INDEX

YU-YUN HSU

Department of Applied Mathematics, National Dong Hwa University
Hualien 974, Taiwan, ROC
yyhsu0607@yahoo.com.tw

SZE-MAN TSE

Department of Applied Mathematics, National Dong Hwa University
Hualien 974, Taiwan, ROC
tse@server.am.ndhu.edu.tw

BERLIN WU*

Department of Mathematical Sciences, National Chengchi University
Wenshan District, Taipei 11623, Taiwan, ROC
berlin@math.nccu.edu.tw

Received February 2002
Revised May 2003

In recent years, the innovation and improvement of forecasting techniques have caught more and more attention. Especially, in the fields of financial economics, management planning and control, forecasting provides indispensable information in decision-making process. If we merely use the time series with the closing price array to build a forecasting model, a question that arises is: Can the model exhibit the real case honestly? Since, the daily closing price of a stock index is uncertain and indistinct. A decision for biased future trend may result in the danger of huge lost. Moreover, there are many factors that influence daily closing price, such as trading volume and exchange rate, and so on. In this research, we propose a new approach for a bivariate fuzzy time series analysis and forecasting through fuzzy relation equations. An empirical study on closing price and trading volume of a bivariate fuzzy time series model for Taiwan Weighted Stock Index is constructed. The performance of linguistic forecasting and the comparison with the bivariate ARMA model are also illustrated.

Keywords: Fuzzy relation; fuzzy Markov relation matrix; bivariate fuzzy time series; fuzzy rule base; mean absolute forecasting accuracy.

1. Introduction

In the humanities and social sciences, fuzzy statistics and fuzzy correlation have gradually received attention. This is a natural result because the complicated phe-

*Tel: 886-2-29387387, Fax: 886-2-29390005

nomenon of humanities and society is hard to be fully explained by traditional models. Taking stock market as an example, the essence of closing price is uncertain and indistinct. Moreover, there are many factors that influence closing price, such as trading volume and exchange rate, etc. Therefore, if we merely consider closing price of yesterday to construct our forecasting model, not only will we misestimate the future trend, but also we will suffer unnecessary loss.

When we apply fuzzy logic in the time series analysis, the first step is to how to integrate linguistic variable analysis methods in solving the autoregressive relation problem of the dynamic data. Graham and Newell (1989), Xu and Lee (1987) presented self-learning methods to modify fuzzy models for dynamic system in linguistic field and later Chiang et. al. (2000) proposed a fuzzy linguistic summary as one of the data mining function to discover useful knowledge from database. In fact, fuzzy relation equations are easier to be understood and applied than decision tables or decision rules.

In view of this, many researchers have adopted fuzzy relation equations for time series analysis and forecasting. For instance, Song and Chissom (1993a, b) proposed the procedure for developing fuzzy time series model by using fuzzy relation equations. Lee et. al. (1994) presented a two-stage fuzzy model identification process combining the linguistic methods with numerical solutions derived from fuzzy relation equations. Wu and Hung (1999) proposed a fuzzy identification procedure for ARCH and Bilinear models. Kumar and Wu (2001) used fuzzy statistical techniques in change period's detection of nonlinear time series. Chen and Hwang (2000) proposed the two-factors time-variant fuzzy time series model and developed two algorithms for temperature prediction. Huarng (2001) proposed heuristic models by integrating problem-specific heuristic knowledge with Chen's (2000) model to improve forecasting.

Yet, their methods did not include the concept of Box-Jenkins' model (ARIMA), such as three steps' construction: order identification, parameters estimation and diagnostic checking. Though Tseng et. al. (2001) proposed a fuzzy ARIMA model which uses the fuzzy regression method to fuzzify the parameters of the ARIMA model. Tseng and Tzeng (2002) combined the SARIMA model and fuzzy regression model to develop the fuzzy SARIMA model. However, literatures in the past have been focusing on univariate fuzzy time series, and it is not easy to apply these techniques into the multivariate systems.

In this research, we propose a bivariate fuzzy time series modeling process. We apply this method to the time series of the closing price and trading volume analysis, and then forecast future trend while comparing the forecasting performance by mean absolute forecasting accuracy. From the empirical studies, it is shown that our proposed method demonstrates an appropriate and efficient performance of prediction for bivariate variables.

2. Bivariate Fuzzy Time Series

2.1. *Fuzzy time series*

The increasing structural complexity of objects which modern scientists desired to study and the fuzziness of human language due to subjectivity, time variation, environmental change, and different angle of research have made scientists difficult hard to clearly investigate the real essence of objects. Let alone to properly construct mathematical models. In response to it, the concept of fuzzy theory is developed upon.

From the angle of economic and social issues, when analyzing human thinking and behavior model, we can often find that the boundary between events is not clear. The fuzzy logic steps in as a viable solution for this problem. The concept of fuzzy logic primarily focuses on individual preferences without the needs of precise and clear numbers. Therefore, it is somewhat different from the essence of Boolean logic. However, because of the complexity of human thinking, language and the ambiguity of preferences, the operation of fuzzy logic is more complicated than Boolean logic. In view of this, it is more suitable to evaluate the correlation between objects by using fuzzy models instead of assigning specific numbers to objects. Though fuzzy logic is a complement of Boolean logic, it is not a replacement. Both theories exist for their reasons. The key point is how to apply these two different thinking models in daily life.

In traditional social and economic researches, they dedicated to the interactive relation and model analysis about human while usually facing some uncertain factors in modeling. For example, should the number of freshman enrollment be counted from the beginning, middle or end of year? Should we determine the exchange rate of NT dollars to US dollars by opening price, closing price, or the average of price ceiling and price floor? The results can vary to a great extent. Hendershot and Placek (1981) had made an extensive review on journals of this field. Moreover, few questions in the social science and economic researches are ultimately true or false. And if we do attempt to analyze human faith, we will soon be facing the great uncertainty behind many behaviors. The consecutive intervals in fuzzy sets have the capability to deal with the gray area between true and false. If this fuzzy characteristic of intervals is applied in analysis, researchers will be able to handle the uncertainty of factors. Indeed, it is a more practical measuring tool in real application.

In this research, we attempt to transform observations into fuzzy sets by using membership functions. Since only through membership functions can we quantify fuzzy sets and further analyze fuzzy information by implementing precise mathematic methods. In order to develop a fuzzy model for observations or estimate fuzzy outputs through fuzzy models, the first is to transform observations into fuzzy sets before tackling fuzzy time series. The fuzzy time series is a method combining linguistic variables with the process of applying fuzzy logic into time series to solve the fuzziness of data.

Thus, before developing bivariate fuzzy time series model and forecasting, we must give some related definitions for fuzzy time series,

Definition 1. (Fuzzy time series) Let $\{X_t \in R, t = 1, 2, \ldots, n\}$ be a time series, Ω be the range of $\{X_t \in R, t = 1, 2, \ldots, n\}$ and $\{P_i; i = 1, 2, \ldots, r, \bigcup_{i=1}^{r} P_i = \Omega\}$ be an ordered partition on Ω. Let $\{L_i, i = 1, 2, \ldots, r\}$ denote linguistic variables with respect to the ordered partition set. For $t = 1, 2, \ldots, n$, if $\mu_i(X_t)$, the grade of membership of X_t belongs to L_i, satisfies $\mu_i : R \longrightarrow [0, 1]$ and $\sum_{i=1}^{r} \mu_i(X_t) = 1$, then $\{FX_t\}$ is said to be a fuzzy time series of $\{X_t\}$ and written as

$$FX_t = \mu_1(X_t)/L_1 + \mu_2(X_t)/L_2 + \cdots + \mu_r(X_t)/L_r,$$

where / is employed to link the linguistic variables with their memberships in FX_t, and the + indicates, rather than any sort of algebraic addition, that the listed pairs of linguistic variables and memberships collectively.

For convenience, let us denote FX_t as $FX_t = (\mu_1, \mu_2, \ldots, \mu_r)$.

When calculating corresponding memberships of linguistic variables in fuzzy time series, this research uses triangular membership functions for facilitating transformation process.

Example 1. Consider the time series $\{X_t\} = \{0.8, 1.7, 2.6, 4.1, 2.9, 3.2, 4.5, 3.8\}$. Let $\Omega = [0, 5]$ and choose an ordered partition set $\{[0, 1), [1, 2), [2, 3), [3, 4), [4, 5]\}$ on Ω. Let $\{L_1, L_2, L_3, L_4, L_5\}$ denote linguistic variables:

$$L_1 \propto [0, 1) : \text{Very low}; L_2 \propto [1, 2) : \text{Low};$$

$$L_3 \propto [2, 3) : \text{Medium}; L_4 \propto [3, 4) : \text{High}; L_5 \propto [4, 5] : \text{Very high}.$$

We evaluate the mean $\{m_1 = 0.5, m_2 = 1.5, m_3 = 2.5, m_4 = 3.5, m_5 = 4.5\}$ of the ordered partition set. Since X_1 is between 0.5 and 1.5, and

$$\frac{1.5 - 0.8}{1.5 - 0.5} = 0.7 \in L_1, \frac{0.8 - 0.5}{1.5 - 0.5} = 0.3 \in L_2,$$

we get the fuzzy set FX_1 with respect to X_1 is $(0.7, 0.3, 0, 0, 0)$. Finally, we can get the fuzzy time series $\{FX_t\}$ of $\{X_t\}$ as follow:

	Very low	Low	Medium	High	Very high	
$FX_1 = ($	0.7,	0.3,	0,	0,	0	)
$FX_2 = ($	0,	0.8,	0.2,	0,	0	)
$FX_3 = ($	0,	0,	0.9,	0.1,	0	)
$FX_4 = ($	0,	0,	0,	0.4,	0.6	)
$FX_5 = ($	0,	0,	0.6,	0.4,	0	)
$FX_6 = ($	0,	0,	0.3,	0.7,	0	)
$FX_7 = ($	0,	0,	0,	0,	1	)
$FX_8 = ($	0,	0,	0,	0.7,	0.3	)

In time series modeling and analysis, the determination of autocorrelation value is very important. But for a set of uncertain or incomplete data, its autocorrelation should not be explained with a number. Therefore, this research attempts to use fuzzy relation to analyze the autocorrelation level in the fuzzy time series.

To begin, we will define general fuzzy relations necessary for exploring bivariate fuzzy time series later on.

Definition 2. (Fuzzy relation) Consider an ordered partition set $\{P_i, i = 1, 2, \dots, r\}$ on Ω. Let $G = (\mu_1, \dots, \mu_r)$ and $H = (\nu_1, \dots, \nu_r)$ be fuzzy sets, then a fuzzy relation R between G and H is

$$R = G^t \circ H = [R_{ij}]_{r \times r},$$

where μ_i, ν_j denote memberships, t denotes transpose, and $R_{ij} = \min(\mu_i, \nu_j)$.

2.2. *Calculation of fuzzy Markov relation matrix* $\mathfrak{R}$

From Section 2.1, we can find that fuzzy relation is the key for constructing good fuzzy time series models. If we can precisely handle fuzzy relation matrix through fuzzy relation, then fuzzy time series models will provide a better fitting result. Besides, there are many different ways for calculating a fuzzy relation matrix. Dubois and Prade (1991), Wu (1986) had proposed some methods to calculate fuzzy relation matrix but none of them is based on the same premises.

Firstly, we consider the stationary bivariate fuzzy time series, and assume that the underlying time series have the Markov property. So, let us define fuzzy Markov relation matrix $\mathfrak{R}$ before constructing bivariate fuzzy time series models,

Definition 3. (Fuzzy Markov relation matrix) The fuzzy time series $\{FX_t, t = 1, 2, \dots, n\}$ is an autoregressive process of order one, FAR(1), that is, FX_t depends only on FX_{t-1}, for all t. Let FX_t has finite memberships $\mu_i(X_t)$, $i = 1, 2, \dots, r$, than the fuzzy Markov relation matrix can be written as

$$\mathfrak{R} = [\mathfrak{R}_{ij}]_{r \times r} = \max_{2 \le t \le n} [\min(\mu_i(X_{t-1}), \mu_j(X_t))]_{r \times r}.$$

Example 2. Consider a fuzzy time series $\{FX_t\}$ of Example 1. Suppose that this fuzzy time series $\{FX_t\}$ is an autoregressive process of order one and detect the linguistic variable according to the position of the greatest membership. We can find the relationships with linguistic variables for this fuzzy time series as follow:

$$L_1 \longrightarrow L_2; L_2 \longrightarrow L_3; L_3 \longrightarrow L_5; L_5 \longrightarrow L_3; L_3 \longrightarrow L_4; L_4 \longrightarrow L_5; L_5 \longrightarrow L_4.$$

Since L_1 and L_2 represent $(1, 0.5, 0, 0, 0)$ and $(0.5, 1, 0.5, 0, 0)$, respectively. By Definition 2, we get the fuzzy relation of $L_1 \longrightarrow L_2$ is

$$R_1 = \begin{bmatrix} 1 \\ 0.5 \\ 0 \\ 0 \\ 0 \end{bmatrix} [0.5 \ 1 \ 0.5 \ 0 \ 0] = \begin{bmatrix} 0.5 & 1 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Similarly, we can get the fuzzy relations of $\{FX_t\}$ as follow:

$$R_2 = \begin{bmatrix} 0 & 0.5 & 0.5 & 0.5 & 0 \\ 0 & 0.5 & 1 & 0.5 & 0 \\ 0 & 0.5 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, R_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0.5 & 1 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, R_4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0.5 & 0 \\ 0 & 0.5 & 1 & 0.5 & 0 \end{bmatrix},$$

$$R_5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 1 & 0.5 \\ 0 & 0 & 0.5 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, R_6 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0.5 & 1 \\ 0 & 0 & 0 & 0.5 & 0.5 \end{bmatrix}, R_7 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 1 & 0.5 \end{bmatrix}.$$

Finally, by Definition 3, we get the following fuzzy Markov relation matrix.

$$\mathfrak{R} = \begin{bmatrix} 0.5 & 1 & 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & 1 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 & 1 & 1 \\ 0 & 0.5 & 0.5 & 0.5 & 1 \\ 0 & 0.5 & 1 & 1 & 0.5 \end{bmatrix}.$$

3. Modeling Bivariate Fuzzy Time Series and Forecasting

3.1. *Modeling bivariate fuzzy time series*

The patterns of collected data can be numerical or qualitative formats, or linguistic values (such as data derived from testing). For these kinds of data, it is hard to analyze by traditional time series models. Therefore, if using fuzzy sets, the patterns of data will not be restricted and a more suitable model can be established.

There is no certain rule for the optimal partition in building fuzzy range sets. Generally, the more partition we do, the more precision we have. However, it also requires more complicated calculation. In short, the determination between accuracy and complexity is entirely up to the individual requirements.

Definition 4. (The Bivariate Fuzzy AR(1) (BFAR(1))) Let $\{(FX_{1,t}, FX_{2,t})\}$ be a bivariate fuzzy time series on the universe domain U. If $\{(FX_{1,t}, FX_{2,t})\}$ can be written as:

$$(FX_{1,t}, FX_{2,t}) = (FX_{1,t-1}, FX_{2,t-1}) \circ \begin{pmatrix} \mathfrak{R}_{11} & \mathfrak{R}_{12} \\ \mathfrak{R}_{21} & \mathfrak{R}_{22} \end{pmatrix},$$

for all t, where $\mathfrak{R}_{ij}$ denote fuzzy Markov relation matrix of $\{FX_{i,t}\}$ and $\{FX_{j,t}\}$, $i, j = 1, 2$. Then we say that $\{(FX_{1,t}, FX_{2,t})\}$ is a bivariate fuzzy autoregressive process of ordered one.

3.2. *Decision support by a fuzzy rule base*

There are many factors influencing stock index trend such as trading volume, exchange rates, interest rates, political causes, and so on. Through fuzzy Markov relation matrix, we can get a decision by transforming the previous fuzzy numbers into current fuzzy numbers operation. The question here is how to transfer fuzzy numbers into corresponding linguistic variables. Firstly, we provide the following definition,

Definition 5. (Linguistic indicator function) Let $(L_{i1}, \ldots, L_{ir})$ be a linguistic vector of $\{FX_i\}$ and $F\widehat{X}_{i,t}$ be a numerical vector of corresponding linguistic vector, $i = 1, 2$. Then $F\widetilde{X}_{i,t} = \{(I_{i1_t}, \ldots, I_{ir_t}); I_{ij_t} = 1 \text{ or } 0; j = 1, \ldots, r\}$ is said to be linguistic indicator function and

$$I_{ij_t} = \begin{cases} 1, & \text{if } \nu_{L_{ij}}(F\widehat{X}_{i,t}) \geq 2 \\ 0, & \text{if } \nu_{L_{ij}}(F\widehat{X}_{i,t}) < 2 \end{cases}, i = 1, 2; j = 1, \ldots, r \text{ and } t = 1, 2, \ldots, n.$$

where $\nu_{L_{ij}}(F\widehat{X}_{i,t})$ denote value of $F\widehat{X}_{i,t}$ in L_{ij}.

Example 3. Let $\{[(L_{11}, L_{12}, L_{13}, L_{14}, L_{15}), (L_{21}, L_{22}, L_{23}, L_{24}, L_{25})]; L_{11}$: plunge, L_{12} : drop, L_{13} : draw, L_{14} : soar, L_{15} : surge; L_{21} : very low, L_{22} : low, L_{23} : medium, L_{24} : high, L_{25} : very high$\}$ be a bivariate linguistic vector of $\{(FX_{1,t}, FX_{2,t})\}$. After calculating a bivariate fuzzy time series data by fuzzy Markov relation matrix, we have $[F\widehat{X}_{1,t}, F\widehat{X}_{2,t}] = [(1.5, 1.5, 1.5, 1.5, 2), (1.5, 1, 1.5, 2, 2)]$. By Definition 5, we get $[F\widetilde{X}_{1,t}, F\widetilde{X}_{2,t}] = [(0, 0, 0, 0, 1), (0, 0, 0, 1, 1)]$.

According to Definition 5, we can transfer fuzzy numbers predicted by bivariate fuzzy time series model into linguistic indicator functions. Yet, the problem is how to determine corresponding linguistic variables through linguistic indicator functions. To solve this, we use Definition 5 and establish a threshold function by fuzzy reasoning to obtain a fuzzy rule base and further analyze its outputting linguistic variables.

According to the times series data used in this research, we build ranges as {plunge, drop, draw, soar, surge} and {very low, low, medium, high, very high} for price limit and trading volume difference and thus set the value of n as 5, respectively. We also use $(I_{i1_t}, \ldots, I_{i5_t})$ as fuzzy inference indicator, where $I_{ij_t} = 1$ or 0 and $j = 1, 2, \ldots, 5$, and thus 32 linguistic indicator functions can be established. Yet, we need to exclude vector $(0, 0, 0, 0, 0)$ because it cannot represent any linguistic variable. However, it's not easy to categorize 31 linguistic indicator functions to their corresponding linguistic variables. If only a "1" appears in the linguistic indicator function, the output will be the corresponding linguistic variable where this 1 is located. For example, $(0, 0, 0, 1, 0)$ represents that the membership of "soar" is "1", so the outputting of linguistic variable is "soar."

However, how to deal with the scenario when there is more than one "1" appearing in the linguistic indicator function will be very time-consuming if we attempt to sequentially examine each component of linguistic indicator function in fuzzy time series, i.e. examining each i_{ij_t}, where $j = 1, 2, \ldots, 5$, from $I_{i1_t} = 1$ or 0 until $I_{i5_t} = 1$ or 0. Yet, if we take entire linguistic indicator function to judge, it will be easy for us to identify its representative linguistic variable through experience rules. For example, $(0, 0, 0, 1, 1)$ represents both the memberships of the linguistic terms "soar" and "surge" are 1. Through experience rules, we can detect its outputting linguistic variable is "surge." Similarly, $(1, 1, 0, 0, 0)$ represents that the outputting linguistic variable is "plunge."

Therefore, using the above methodology, we consider the threshold function H_t as our decision process for different range partition sets, where H_t is defined as follows:

$$H_t = \begin{cases} \text{plunge (very low)} & \text{, if } K_t \leq -2 \\ \text{drop (low)} & \text{, if } -2 < K_t \leq -1 \text{ ; or if } K_t = -2 \text{ and } \sum_{j=1}^{5} I_{ij_t} \geq 3 \\ \text{draw (medium)} & \text{, if } K_t = 0 \\ \text{soar (high)} & \text{, if } 1 \leq K_t < 2 \quad \text{; or if } K_t = 2 \text{ and } \sum_{j=1}^{5} I_{ij_t} \geq 3 \\ \text{surge (very high)} & \text{, if } 2 \leq K_t \end{cases},$$

where $K_t = \sum_{j=1}^{5} (j - 3) I_{ij_t}$, $i = 1, 2$.

Finally, we can use this threshold function to establish the following fuzzy rule base.

A fuzzy rule base

For $i = 1, 2$.

(i) If $F\widetilde{X}_{i,t} \in \{(1,0,0,0,0), (1,1,0,0,0), (1,0,1,0,0), (1,1,1,0,0)\}$, then the outputting linguistic variable is "plunge (very low)".

(ii) If $F\widetilde{X}_{i,t} \in \{(0,1,0,0,0), (1,1,0,1,0), (1,1,1,0,1), (1,1,0,0,1), (1,0,0,1,0), (1,1,1,1,0), (0,1,1,0,0), (1,0,1,1,0)\}$, then the outputting linguistic variable is "drop (low)".

(iii) If $F\widetilde{X}_{i,t} \in \{(0,0,1,0,0), (1,0,1,0,1), (1,0,0,0,1), (1,1,1,1,1), (0,1,0,1,0), (1,1,0,1,1), (0,1,1,1,0)\}$, then the outputting linguistic variable is "draw (medium)".

(iv) If $F\widetilde{X}_{i,t} \in \{(0,0,0,1,0), (0,1,0,1,1), (1,0,1,1,1), (1,0,0,1,1), (0,1,0,0,1), (0,1,1,1,1), (0,0,1,1,0), (0,1,1,0,1)\}$, then the outputting linguistic variable is "soar (high)".

(v) If $F\widetilde{X}_{i,t} \in \{(0,0,0,0,1), (0,0,0,1,1), (0,0,1,0,1), (0,0,1,1,1)\}$, then the outputting linguistic variable is "surge (very high)".

For instance, in Example 3, we obtained $[F\widetilde{X}_{1,t}, F\widetilde{X}_{2,t}] = [(0,0,0,0,1), (0,0,0,1,1)]$. Therefore, through fuzzy rule base mentioned above, we can get the outputting linguistic variables for price limit and trading volume difference which are "surge" and "very high", respectively.

3.3. *Mean absolute forecasting accuracy*

Forecasting provides indispensable information in decision-making process. Especially, a precise forecasting result can provide decision makers precious information to make correct decision and appropriate reaction. To which, we use bivariate fuzzy times series model for forecasting to realize its predictive effects. After modeling the bivariate fuzzy time series and constructing the fuzzy rule base, it is necessary to compare the forecasting performance with other methods.

For the bivariate fuzzy autoregressive process of ordered one, the l-step prediction becomes

$$(FX_{1,n}(l), FX_{2,n}(l)) = (FX_{1,n}, FX_{2,n}) \circ \begin{pmatrix} \mathfrak{R}_{11} & \mathfrak{R}_{12} \\ \mathfrak{R}_{21} & \mathfrak{R}_{22} \end{pmatrix}^{l}.$$

To compare the forecasting performances, we need to assign each linguistic variable with an ordered rank. For instance, a "plunge" as -2, "drop" as -1, "draw" as 0, "soar" as 1 and "surge" as 2. By doing so, the mean absolute forecasting accuracy can be defined.

Definition 6. (Mean Absolute Forecasting Accuracy) Suppose $\{RL_t, t = 1, \ldots, n\}$ and $\{FL_t, t = 1, \ldots, n\}$ denote respectively the real and outputting linguistic variables. Let $L = \{(L_1, L_2, \ldots, L_r) = (-(r-1)/2, -(r-3)/2, \ldots, (r-1)/2; L_j$: linguistic variable, $j = 1, \ldots, r\}$ be corresponding values of linguistic variables, then $MAFA$ is said to be mean absolute forecasting accuracy and written as

$$MAFA = 1 - \frac{\sum_{t=1}^{n} \frac{|FL_t - RL_t|}{r-1}}{n}.$$

where r denote the number of linguistic variables.

Example 4. Suppose that real linguistic variables of the time series are {drop, draw, drop, surge, draw, drop, surge, drop, draw, plunge}, then the corresponding values of linguistic variables are $\{-1, 0, -1, 2, 0, -1, 2, -1, 0, -2\}$. The outputting linguistic variables are {drop, draw, plunge, surge, draw, draw, surge, drop, surge, draw}, then the corresponding values of linguistic variables are $\{-1, 0, -2, 2, 0, 0, 2, -1, 2, 0\}$. By Definition 3.3, we can get

$$MAFA = 1 - \frac{\sum_{t=1}^{10} \frac{|FL_t - RL_t|}{4}}{10} = 1 - \frac{6}{40} = 0.85.$$

We further provide the integrated process for bivariate fuzzy time series modeling,

An algorithm for modeling BFAR(1)

Step 1. For time series $\{X_{1,t}\}, \{X_{2,t}\}$. Decide the range of Ω_i and the linguistic variables $\{L_{i1}, L_{i2}, \ldots, L_{ir}\}$ of $\{X_{i,t}\}$, $i = 1, 2$.

Step 2. Calculate the fuzzy time series $\{FX_{i,t}\}$ of $\{X_{i,t}\}$, $i = 1, 2$ and detect the linguistic variable according to the position of the greatest membership in $\{FX_{i,t}, t = 1, 2, \ldots, n\}$, $i = 1, 2$.

Step 3. Calculate all fuzzy relations between $\{FX_{i,t}\}$ and $\{FX_{j,t}\}$, $i, j = 1, 2$ to get the fuzzy Markov relation matrix, then construct a bivariate fuzzy time series model.

Step 4. Examine $\{F\widetilde{X}_{i,t}, t = 1, 2, \ldots, n\}$, $i = 1, 2$. If the number of "1" is only one, we can detect the corresponding linguistic variable immediately, otherwise detect the corresponding linguistic variables by fuzzy rule base.

Step 5. Forecast by bivariate fuzzy time series model.

Step 6. Stop.

4. An Empirical Application

4.1. *Data analysis*

These data source comes from Taiwan Stock Exchange Corporation, including daily price limit and trading volume difference of weighted index from December 30 2000 to February 9 2001. The tendencies of these data are shown respectively in Fig. 1 and Fig. 2.

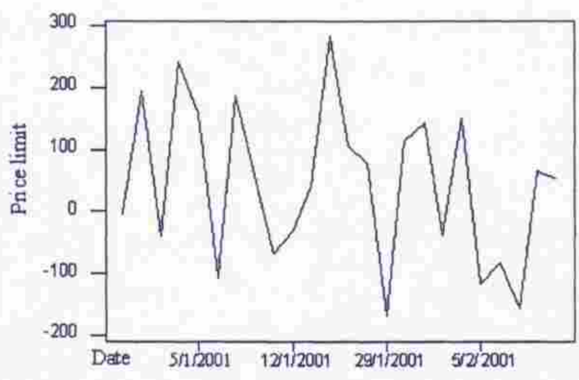

Fig. 1. Trend for price limit of Taiwan Weighted Stock Index

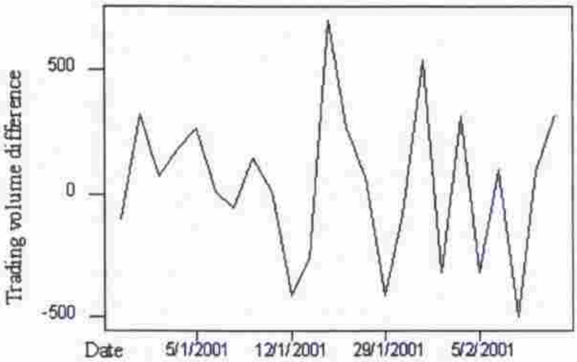

Fig. 2. Trend for trading volume difference of Taiwan Weighted Stock Index

From these data, we can get that maximum price limit of weighted index is 283.28 and minimum one is -167.85. Also, maximum trading volume difference of weighted index is 699 and minimum one is -496. Generally speaking, range should include maximum and minimum so we take two sets, $(-167.85, 283.28)$ and $(-496, 699)$ as the ranges of daily price limit and trading volume difference of weighted index, respectively. Because this research is based on fuzzy theory, we must fuzzify data so that modeling process can be started. Therefore, we first partition $(-167.85, 283.28)$ and $(-496, 699)$ into k intervals (we set $k = 5$ here), respectively, i.e.

$I_{11} = (X_{(1)}, (X_{(3)} + X_{(4)})/2) = (-167.85, -111.93)$,
the representative value of I_{11} is -167.85;
$I_{12} = ((X_{(3)} + X_{(4)})/2, (X_{(9)} + X_{(10)})/2) = (-111.93, -17.345)$,
the representative value of I_{12} is -54.015;
$I_{13} = ((X_{(9)} + X_{(10)})/2, (X_{(15)} + X_{(16)})/2) = (-17.345, 92.485)$,
the representative value of I_{13} is 55.95;
$I_{14} = ((X_{(15)} + X_{(16)})/2, (X_{(21)} + X_{(22)})/2) = (92.485, 191.9)$,
the representative value of I_{14} is 147.515;
$I_{15} = ((X_{(21)} + X_{(22)})/2, X_{(24)}) = (191.9, 283.28)$,
the representative value of I_{15} is 283.28.

$I_{21} = (X_{(1)}, (X_{(3)} + X_{(4)})/2) = (-496, -365.5)$,
the representative value of I_{21} is -496;
$I_{22} = ((X_{(3)} + X_{(4)})/2, (X_{(9)} + X_{(10)})/2) = (-365.5, -22)$,
the representative value of I_{22} is -183;
$I_{23} = ((X_{(9)} + X_{(10)})/2, (X_{(15)} + X_{(16)})/2) = (-22, 121)$,
the representative value of I_{23} is 66.5;
$I_{24} = ((X_{(15)} + X_{(16)})/2, (X_{(21)} + X_{(22)})/2) = (121, 316.5)$,
the representative value of I_{24} is 265.5;
$I_{25} = ((X_{(21)} + X_{(22)})/2, X_{(24)}) = (316.5, 699)$,
the representative value of I_{25} is 699.

where $\{I_{11}, I_{12}, I_{13}, I_{14}, I_{15}\}$ and $\{I_{21}, I_{22}, I_{23}, I_{24}, I_{25}\}$ are five intervals of $(-167.85, 283.28)$ and $(-496, 699)$, respectively. We also define five linguistic variables with respect to the above intervals for $(-167.85, 283.28)$ and $(-496, 699)$, respectively, i.e.

$L_{11} \propto I_{11}$: "plunge";	$L_{21} \propto I_{21}$: "very low";
$L_{12} \propto I_{12}$: "drop";	$L_{22} \propto I_{22}$: "low";
$L_{13} \propto I_{13}$: "draw";	$L_{23} \propto I_{23}$: "medium";
$L_{14} \propto I_{14}$: "soar";	$L_{24} \propto I_{24}$: "high";
$L_{15} \propto I_{15}$: "surge".	$L_{25} \propto I_{25}$: "very high".

where each linguistic variable denotes a fuzzy set and each element in fuzzy set denotes by I_{ij} $(i = 1, 2; j = 1, 2, \ldots, 5)$ and its corresponding membership. Therefore, these linguistic variables can be expressed by fuzzy set as follows,

$$\begin{aligned}
L_{i1} &= \{1/I_{i1}, 0.5/I_{i2}, 0/I_{i3}, 0/I_{i4}, 0/I_{i5}\} \\
L_{i2} &= \{0.5/I_{i1}, 1/I_{i2}, 0.5/I_{i3}, 0/I_{i4}, 0/I_{i5}\} \\
L_{i3} &= \{0/I_{i1}, 0.5/I_{i2}, 1/I_{i3}, 0.5/I_{i4}, 0/I_{i5}\}, \quad i = 1, 2. \\
L_{i4} &= \{0/I_{i1}, 0/I_{i2}, 0.5/I_{i3}, 1/I_{i4}, 0.5/I_{i5}\} \\
L_{i5} &= \{0/I_{i1}, 0/I_{i2}, 0/I_{i3}, 0.5/I_{i4}, 1/I_{i5}\}
\end{aligned}$$

where the determination of memberships is based on principle proposed by Song and Chissom (1993). For convenience, the elements in fuzzy set can be expressed by their corresponding memberships.

$$\begin{aligned}
L_{i1} &= (1, 0.5, 0, 0, 0) \\
L_{i1} &= (0.5, 1, 0.5, 0, 0) \\
L_{i1} &= (0, 0.5, 1, 0.5, 0), \quad i = 1, 2. \\
L_{i1} &= (0, 0, 0.5, 1, 0.5) \\
L_{i1} &= (0, 0, 0, 0.5, 1)
\end{aligned}$$

4.2. *Fuzzy model construction*

After fuzzifying these data of daily price limit and trading volume difference of weighted index, we can apply the method mentioned in Section 2.1 to calculate data's corresponding memberships in L_{ij} $(i = 1, 2; j = 1, 2, \ldots, 5)$. We illustrate the results in Table 1 and Table 2.

Table 1. The memberships for price limit of Taiwan Weighted Stock Index

Date	Price limit	L_{11}	L_{12}	L_{13}	L_{14}	L_{15}
2000-12-30	−4.85	0	0.42	0.58	0	0
2001-1-2	196.19	0	0	0	0.43	0.57
2001-1-3	−40.49	0	0.76	0.24	0	0
2001-1-4	241.34	0	0	0	0	1
2001-1-5	159.4	0	0	0	0.82	0.18
2001-1-8	−107.02	0.56	0.44	0	0	0
2001-1-9	187.61	0	0	0	0.52	0.48
2001-1-10	60.66	0	0	0.78	0.22	0
2001-1-11	−67.54	0.04	0.96	0	0	0
2001-1-12	−29.84	0	0.66	0.34	0	0
2001-1-15	40.26	0	0	0.97	0.03	0
2001-1-16	283.28	0	0	0	0	1
2001-1-17	107.01	0	0	0.34	0.66	0
2001-1-18	77.96	0	0	0.61	0.39	0
2001-1-29	−167.85	1	0	0	0	0
2001-1-30	112.44	0	0	0.28	0.72	0
2001-1-31	143.7	0	0	0	0.98	0.02
2001-2-1	−38.27	0	0.74	0.26	0	0
2001-2-2	151.33	0	0	0	0.9	0.1
2001-2-5	−116.84	0.69	0.31	0	0	0
2001-2-6	−83.36	0.25	0.75	0	0	0
2001-2-7	−155.48	1	0	0	0	0
2001-2-8	65.02	0	0	0.74	0.26	0
2001-2-9	51.24	0	0	0.87	0.13	0

Table 2. The memberships for trading volume difference of Taiwan Weighted Stock Index

Date	Trading volume difference	L_{21}	L_{22}	L_{23}	L_{24}	L_{25}
2000-12-30	−102	0	0.62	0.38	0	0
2001-1-2	317	0	0	0	0.66	0.34
2001-1-3	70	0	0	0.88	0.12	0
2001-1-4	175	0	0	0.26	0.74	0
2001-1-5	263	0	0	0	0.85	0.15
2001-1-8	12	0	0.15	0.85	0	0
2001-1-9	−56	0	0.43	0.57	0	0
2001-1-10	146	0	0	0.43	0.57	0
2001-1-11	12	0	0.15	0.85	0	0
2001-1-12	−412	0.92	0.08	0	0	0
2001-1-15	−264	0.3	0.7	0	0	0
2001-1-16	699	0	0	0	0	1
2001-1-17	268	0	0	0	0.83	0.17
2001-1-18	63	0	0	0.92	0.08	0
2001-1-29	−413	0.93	0.07	0	0	0
2001-1-30	−82	0	0.54	0.46	0	0
2001-1-31	536	0	0	0	0	1
2001-2-1	−319	0.53	0.47	0	0	0
2001-2-2	313	0	0	0	0.67	0.33
2001-2-5	−319	0.53	0.47	0	0	0
2001-2-6	96	0	0	0.73	0.27	0
2001-2-7	−496	1	0	0	0	0
2001-2-8	95	0	0	0.73	0.27	0
2001-2-9	316	0	0	0	0.66	0.34

From Table 1 and Table 2, we can see the memberships of daily price limit and trading volume difference of weighted index from December 30 2000 to February 9 2001. If the day's greatest membership of price limit is fallen within L_{1j} ($j = 1, 2, \ldots, 5$), we can consider this day's price limit belonging to this linguistic variable L_{1j} ($j = 1, 2, \ldots, 5$). Taking data of December 30 2000 as an example, where the greatest membership is fallen within L_{13} and L_{22}, we can consider this day's price limit and trading volume difference belonging to L_{13} and L_{22}, respectively. In other words, the price limit is "draw" and trading volume difference is "low". From past fuzzy data, we can find fuzzy relations between data and further obtain fuzzy Markov relation matrix.

Therefore, we can explore the fuzzy cross correlation for two consecutive data, price limit and trading volume difference. If the price limit and trading volume difference on date $(t-1)$ is $(L_{1m}(t-1), L_{2u}(t-1))$ and the trading volume difference on date (t) is $(L_{1n}(t), L_{2v}(t))$, the fuzzy relations between date $(t-1)$ price limit and date (t) respectively are $L^t_{1m} \times L_{1n}$, $L^t_{1m} \times L_{2v}$, $L^t_{2u} \times L_{1n}$, and $L^t_{2u} \times L_{2v}$ (the subscript for time is omitted here). Hence, we can get fuzzy relations as follows,

(Price limit $\longrightarrow$ Price limit)
$R_1 = L_{13}^t \times L_{25}$; $R_2 = L_{15}^t \times L_{22}$; $R_3 = L_{12}^t \times L_{25}$; $R_4 = L_{15}^t \times L_{24}$;
$R_5 = L_{14}^t \times L_{21}$; $R_6 = L_{11}^t \times L_{24}$; $R_7 = L_{14}^t \times L_{23}$; $R_8 = L_{13}^t \times L_{22}$;
$R_9 = L_{12}^t \times L_{22}$; $R_{10} = L_{12}^t \times L_{23}$; $R_{11} = L_{13}^t \times L_{21}$; $R_{12} = L_{14}^t \times L_{24}$;
$R_{13} = L_{14}^t \times L_{22}$; $R_{14} = L_{12}^t \times L_{24}$; $R_{15} = L_{11}^t \times L_{22}$; $R_{16} = L_{12}^t \times L_{21}$;
$R_{17} = L_{11}^t \times L_{23}$; $R_{18} = L_{13}^t \times L_{23}$.

(Price limit $\longrightarrow$ Trading volume difference)
$R_1 = L_{13}^t \times L_{24}$; $R_2 = L_{15}^t \times L_{23}$; $R_3 = L_{12}^t \times L_{24}$; $R_4 = L_{15}^t \times L_{24}$;
$R_5 = L_{14}^t \times L_{23}$; $R_6 = L_{11}^t \times L_{23}$; $R_7 = L_{14}^t \times L_{24}$; $R_8 = L_{13}^t \times L_{23}$;
$R_9 = L_{12}^t \times L_{21}$; $R_{10} = L_{12}^t \times L_{22}$; $R_{11} = L_{13}^t \times L_{25}$; $R_{12} = L_{13}^t \times L_{21}$;
$R_{13} = L_{11}^t \times L_{22}$; $R_{14} = L_{14}^t \times L_{25}$; $R_{15} = L_{14}^t \times L_{21}$.

(Trading volume difference $\longrightarrow$ Price limit)
$R_1 = L_{12}^t \times L_{25}$; $R_2 = L_{14}^t \times L_{22}$; $R_3 = L_{13}^t \times L_{25}$; $R_4 = L_{14}^t \times L_{24}$;
$R_5 = L_{14}^t \times L_{21}$; $R_6 = L_{13}^t \times L_{24}$; $R_7 = L_{13}^t \times L_{23}$; $R_8 = L_{13}^t \times L_{22}$;
$R_9 = L_{11}^t \times L_{23}$; $R_{10} = L_{15}^t \times L_{24}$; $R_{11} = L_{14}^t \times L_{23}$; $R_{12} = L_{13}^t \times L_{21}$;
$R_{13} = L_{11}^t \times L_{24}$; $R_{14} = L_{12}^t \times L_{24}$; $R_{15} = L_{15}^t \times L_{22}$; $R_{16} = L_{11}^t \times L_{22}$.

(Trading volume difference $\longrightarrow$ Trading volume difference)
$R_1 = L_{12}^t \times L_{24}$; $R_2 = L_{14}^t \times L_{23}$; $R_3 = L_{13}^t \times L_{24}$; $R_4 = L_{14}^t \times L_{24}$;
$R_5 = L_{13}^t \times L_{23}$; $R_6 = L_{13}^t \times L_{21}$; $R_7 = L_{11}^t \times L_{22}$; $R_8 = L_{12}^t \times L_{25}$;
$R_9 = L_{15}^t \times L_{24}$; $R_{10} = L_{15}^t \times L_{21}$; $R_{11} = L_{11}^t \times L_{24}$; $R_{12} = L_{14}^t \times L_{21}$;
$R_{13} = L_{11}^t \times L_{23}$.

where $\times$ denotes minimum operator and t denotes transpose. Taking maximum operator for each part, we can get, $\mathfrak{R}_{11}, \mathfrak{R}_{12}, \mathfrak{R}_{21}, \mathfrak{R}_{22}$, where $\mathfrak{R}_{11}$ is a fuzzy Markov relation matrix for date $(t-1)$ price limit and date (t) price limit, $\mathfrak{R}_{12}$ is a fuzzy Markov relation matrix for date $(t-1)$ price limit and date (t) trading volume difference, $\mathfrak{R}_{21}$ a is fuzzy Markov relation matrix for date $(t-1)$ trading volume difference and date (t) price limit, and $\mathfrak{R}_{22}$ is a fuzzy Markov relation matrix for date $(t-1)$ trading volume difference and date (t) trading volume difference. Finally, the fuzzy Markov relation matrix $\mathfrak{R}$ will be

$$\mathfrak{R} = \begin{bmatrix} \mathfrak{R}_{11} & \mathfrak{R}_{12} \\ \mathfrak{R}_{21} & \mathfrak{R}_{22} \end{bmatrix} = \begin{bmatrix} 0.5 & 1 & 1 & 1 & 0.5 & 0.5 & 1 & 1 & 0.5 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0.5 & 1 & 0.5 \\ 1 & 1 & 1 & 0.5 & 1 & 1 & 0.5 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0.5 & 1 & 0.5 & 1 & 1 & 1 \\ 0.5 & 1 & 0.5 & 1 & 0.5 & 0.5 & 0.5 & 1 & 1 & 0.5 \\ 0.5 & 1 & 1 & 1 & 0.5 & 0.5 & 1 & 1 & 1 & 0.5 \\ 0.5 & 0.5 & 0.5 & 1 & 1 & 0.5 & 0.5 & 0.5 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0.5 & 1 & 1 & 0.5 \\ 1 & 1 & 1 & 1 & 0.5 & 1 & 0.5 & 1 & 1 & 0.5 \\ 0.5 & 1 & 0.5 & 1 & 0.5 & 1 & 0.5 & 0.5 & 1 & 0.5 \end{bmatrix}.$$

Therefore, the bivariate fuzzy autoregressive time series of order one model is $(FX_{1,t}, FX_{2,t}) = (FX_{1,t-1}, FX_{2,t-1}) \circ \mathfrak{R}$, where $(FX_{1,t-1}, FX_{2,t-1})$ and $(FX_{1,t}, FX_{2,t})$ denote bivariate fuzzy sets of price limit and trading volume difference for date $(t-1)$ and date (t), respectively. Finally, we calculate the values from modeling output and values transformed from linguistic indicator function, which is shown at Table 3.

Table 3. The values of modeling and transformed outputs for price limit and trading volume difference of Taiwan Weighted Stock Index

Date	Modeling output	Transformed output
2001-1-2	[(1.5, 1.5, 1.5, 1.5, 2), (1.5, 1, 1.5, 2, 2)]	[(0, 0, 0, 0, 1), (0, 0, 0, 1, 1)]
2001-1-3	[(1.5, 2, 1.5, 2, 1), (1.5, 1, 2, 2, 1)]	[(0, 1, 0, 1, 0), (0, 0, 1, 1, 0)]
2001-1-4	[(2, 2, 2, 2, 2), (2, 1.5, 1.5, 2, 1)]	[(1, 1, 1, 1, 1), (1, 0, 0, 1, 0)]
2001-1-5	[(1.5, 2, 1.5, 2, 1), (1.5, 1, 2, 2, 1)]	[(0, 1, 0, 1, 0), (0, 0, 1, 1, 0)]
2001-1-8	[(2, 2, 2, 2, 1), (2, 1, 2, 2, 1.5)]	[(1, 1, 1, 1, 0), (1, 0, 1, 1, 0)]
2001-1-9	[(1.5, 2, 2, 2, 1.5), (1.5, 1.5, 2, 1.5, 0.5)]	[(0, 1, 1, 1, 0), (0, 0, 1, 0, 0)]
2001-1-10	[(2, 2, 2, 2, 1.5), (2, 1, 2, 2, 1.5)]	[(1, 1, 1, 1, 0), (1, 0, 1, 1, 0)]
2001-1-11	[(2, 2, 2, 1.5, 1.5), (2, 1, 2, 2, 1.5)]	[(1, 1, 1, 0, 0), (1, 0, 1, 1, 0)]
2001-1-12	[(2, 2, 2, 2, 2), (2, 1.5, 1.5, 2, 1)]	[(1, 1, 1, 1, 1), (1, 0, 0, 1, 0)]
2001-1-15	[(1.5, 2, 2, 2, 1.5), (1.5, 2, 1.5, 2, 1)]	[(0, 1, 1, 1, 0), (0, 1, 0, 1, 0)]
2001-1-16	[(1.5, 1.5, 1.5, 1.5, 2), (1.5, 1, 1.5, 2, 2)]	[(0, 0, 0, 0, 1), (0, 0, 0, 1, 1)]
2001-1-17	[(1.5, 2, 1.5, 2, 1), (1.5, 1, 2, 2, 1)]	[(0, 1, 0, 1, 0), (0, 0, 1, 1, 0)]
2001-1-18	[(2, 2, 2, 2, 1), (2, 1, 2, 2, 1.5)]	[(1, 1, 1, 1, 0), (1, 0, 1, 1, 0)]
2001-1-29	[(2, 2, 2, 1.5, 2), (2, 1, 2, 2, 1.5)]	[(1, 1, 1, 0, 1), (1, 0, 1, 1, 0)]
2001-1-30	[(1, 2, 2, 2, 1), (1, 2, 2, 1.5, 0.5)]	[(0, 1, 1, 1, 0), (0, 1, 1, 0, 0)]
2001-1-31	[(1.5, 1.5, 1.5, 2, 1.5), (1.5, 1, 1.5, 2, 2)]	[(0, 0, 0, 1, 0), (0, 0, 0, 1, 1)]
2001-2-1	[(1.5, 2, 1.5, 2, 1), (2, 1, 1.5, 2, 1.5)]	[(0, 1, 0, 1, 0), (1, 0, 0, 1, 0)]
2001-2-2	[(1.5, 2, 2, 2, 1.5), (1.5, 2, 1.5, 2, 1)]	[(0, 1, 1, 1, 0), (0, 1, 0, 1, 0)]
2001-2-5	[(2, 2, 2, 2, 1), (2, 1, 2, 2, 1.5)]	[(1, 1, 1, 1, 0), (1, 0, 1, 1, 0)]
2001-2-6	[(1, 2, 2, 2, 1), (1, 2, 2, 1.5, 0.5)]	[(0, 1, 1, 1, 0), (0, 1, 1, 0, 0)]
2001-2-7	[(2, 2, 2, 2, 2), (2, 1.5, 1.5, 2, 1)]	[(1, 1, 1, 1, 1), (1, 0, 0, 1, 0)]
2001-2-8	[(1, 2, 2, 2, 1), (1, 2, 2, 1.5, 0.5)]	[(0, 1, 1, 1, 0), (0, 1, 1, 0, 0)]
2001-2-9	[(2, 2, 2, 1.5, 2), (2, 1, 2, 2, 1.5)]	[(1, 1, 1, 0, 1), (1, 0, 1, 1, 0)]

4.3. *Forecasting performance*

Because this research is to explore the qualitative trend of time series, we computed the transformed memberships through fuzzy rule base in fuzzy systems for getting their corresponding linguistic variables to facilitate analysis. We already comprehensively defined and introduced fuzzy rule base in Section 3.2 and compared it with bivariate ARMA model (BARMA(p, q)) usually used in analyzing bivariate time series data. The results derived from above principles are shown at Table 4 and Table 5.

Table 4. The comparison of fitted value for price limit of Taiwan Weighted Stock Index

Date	Real value	BARMA(0,1)	FVAR(1)
2001-1-2	Surge	Soar	Surge
2001-1-3	Drop	Surge	Draw
2001-1-4	Surge	Draw	Draw
2001-1-5	Soar	Surge	Draw
2001-1-8	Drop	Plunge	Drop
2001-1-9	Soar	Surge	Draw
2001-1-10	Draw	Drop	Drop
2001-1-11	Drop	Draw	Plunge
2001-1-12	Drop	Draw	Draw
2001-1-15	Draw	Draw	Draw
2001-1-16	Surge	Draw	Surge
2001-1-17	Soar	Draw	Draw
2001-1-18	Draw	Draw	Drop
2001-1-29	Plunge	Draw	Drop
2001-1-30	Soar	Draw	Draw
2001-1-31	Soar	Draw	Soar
2001-2-1	Drop	Surge	Draw
2001-2-2	Soar	Draw	Draw
2001-2-5	Plunge	Draw	Drop
2001-2-6	Drop	Drop	Draw
2001-2-7	Plunge	Drop	Draw
2001-2-8	Draw	Draw	Draw
2001-2-9	Draw	Plunge	Drop
Right		0.17	0.26
Accuracy		0.71	0.78

Table 5. The comparison of fitted value for trading volume difference of Taiwan Weighted Stock Index

Date	Real value	BARMA(0,1)	FVAR(1)
2001-1-2	Very high	High	Very high
2001-1-3	Medium	Medium	High
2001-1-4	High	Low	Low
2001-1-5	High	Very low	High
2001-1-8	Medium	Very high	Low
2001-1-9	Low	Low	Medium
2001-1-10	High	Very high	Low
2001-1-11	Medium	High	Low
2001-1-12	Very low	Low	Low
2001-1-15	Low	Very low	Medium
2001-1-16	Very high	Very high	Very high
2001-1-17	High	Very low	High
2001-1-18	Medium	Very high	Low
2001-1-29	Very low	High	Low
2001-1-30	Low	Medium	Low
2001-1-31	Very high	Low	Very high
2001-2-1	Low	Low	Low
2001-2-2	High	Very high	Medium
2001-2-5	Low	Medium	Low
2001-2-6	Medium	Low	Low
2001-2-7	Very low	Low	Low
2001-2-8	Medium	Medium	Low
2001-2-9	High	Medium	Low
Right		0.22	0.35
Accuracy		0.68	0.8

For comparing the forecasting results, we define mean absolute forecasting accuracy to measure the accuracy of forecasting method. From Table 4 and Table 5, we can find the FVAR(1) model has better forecasting performance than bivariate ARMA model, BARMA(0,1), which indicates that FVAR(1) model is an effective forecasting tool. The prediction for price limit and trading volume difference of weighted index in future four periods are shown at Table 6 and Table 7.

Table 6. The comparison of real and predictive values for price limit of Taiwan Weighted Stock Index

Date	Real value	BARMA(0,1)	FVAR(1)
2001-2-12	Draw	Surge	Plunge
2001-2-13	Soar	Draw	Surge
2001-2-14	Plunge	Draw	Drop
2001-2-15	Surge	Surge	Surge

Table 7. The comparison of real and predictive values for trading volume difference of Taiwan Weighted Stock Index

Date	Real value	BARMA(0,1)	FVAR(1)
2001-2-12	Low	Medium	Low
2001-2-13	Very high	Low	Very high
2001-2-14	High	Low	Low
2001-2-15	Low	Low	Medium

From Table 6 and Table 7, we can get the mean absolute forecasting accuracy for price limit and trading volume difference of weighted index are 0.75 and 0.81, respectively. This result illustrates that the bivariate fuzzy time series model in this research has better forecasting performance. The major reason why the prediction cannot hit real value is that we only consider the greatest membership and omit others memberships. Therefore, only with reasonable forecasting model can we decide investment strategy from forecasting results. Otherwise, without the direction of clear outlines, investors will face a plight as to which information they should take.

5. Conclusion

In this research, we tried to make an appropriate process of constructing bivariate fuzzy time series model and use this model to forecast the price limit and trading volume difference of Taiwan Weighted Stock Index. Compare the bivariate fuzzy time series model with traditional bivariate ARMA model by the performance of mean absolute forecasting accuracy, we can find that the bivariate fuzzy time series model has better forecasting performance than that of traditional bivariate ARMA model. We hope this method will provide a new forecasting technique for investors to make optimal decision with fuzzy information.

Finally, in spite of the forecasting performance for bivariate fuzzy time series modeling, there are some problems for further studies. For example:

(i) To make a general rule for fuzzy order identification instead of the Markov relation?

(ii) To extend our result to the multivariate fuzzy time series case. In fact, how to solve the nonstationary or seasonal factors in the bivariate fuzzy time series are still open questions.

(iii) In this research, we adopt five-ranking classification and transform the time series data into fuzzy numbers through membership functions. However, seven-ranking classification used in social sciences may be used in future studies for special situation. And it is yet to prove where it will provide significant improvement on forecasting performance?

Acknowledgements

The authors are grateful to the referees for their valuable suggestions on the earlier version of this paper.

References

1. B. P. Graham and R. B. Newell, "Fuzzy adaptive control of a first-order process", *Fuzzy Sets and Systems*, **31** (1989) 47–65.
2. C. W. Xu and Y. Z. Lee, "Fuzzy Model Identification and Self-learning for Dynamic Systems", *IEEE Transactions on Systems, Man, and Cybernetics, Vol. SCM-17*, (1987) 683–689.
3. D. Chiang, L. Chow, and Y. Wang, "Mining time series data by a fuzzy linguistic summary system", *Fuzzy Sets and Systems*, **112** (2000) 419–432.
4. Q. Song and B. S. Chissom, "Forecasting enrollments with fuzzy time series - Part I", *Fuzzy Sets and Systems*, **54** (1993a) 1–9.
5. Q. Song and B. S. Chissom, "Fuzzy time series and its models", *Fuzzy Sets and Systems*, **54** (1993b) 269–277.
6. Y. C. Lee, C. Hwnag, and Y. P. Shih, "A combined approach to fuzzy model identification", *IEEE Transactions on Systems, Man, and Cybernetics, Vol. 24*, No. 5 (1994) 736–743.
7. W. Wu, "Fuzzy reasoning and fuzzy relational equations", *Fuzzy Sets and Systems*, **20** (1986) 67–78.
8. K. Kumar and B. Wu, "Detection of change points in time series analysis with fuzzy statistics", *International Journal of Systems Science*, **32**(9) (2001) 1185–1192.
9. S. M. Chen and J. R. Hwang, "Temperature prediction using fuzzy time series", *IEEE Transactions on Systems, Man, and Cybernetics, Vol. 30*, No. 2 (2000) 263–275.
10. K. Huarng, "Heuristic models of fuzzy time series for forecasting", *Fuzzy Sets and Systems*, **123** (2001) 369–386.
11. G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, San Francisco: Holden-Day, (1976).
12. F. Tseng, G. Tzeng, H. Yu, and B. Yuan, *Fuzzy ARIMA model for forecasting the foreign exchange market*, *Fuzzy Sets and Systems*, **118** (2001) 9–19.
13. F. Tseng and G. Tzeng, "A fuzzy SARIMA model for forecasting", *Fuzzy Sets and Systems*, **126** (2002) 367–376.
14. G. Hendershot and P. Placek (Eds), *Predicting Fertility*, Lexington, MA: health and Co. (1981).
15. D. Dubois and H. Prade, "Fuzzy sets in approximate reasoning, Part I: Inference with possibility distributions", *Fuzzy Sets and Systems*, **40** (1991) 143–202.
16. B. Wu and S. Hung, "A fuzzy identification procedure for nonlinear time series: with example on ARCH and bilinear models", *Fuzzy Sets and Systems*, **108** (1999) 175–287.