Review of Economics & Finance
Submitted on 27/October/2011
Article ID: 1923-7529-2012-01-38-13 Tai-Liang Chen

Forecasting the Taiwan Stock Market with a Novel Momentum-based Fuzzy Time-series

Tai-Liang Chen
Dept. of Information Management and Communication,
Wenzao Ursuline College of Languages
900 Mintsu 1st Road Kaohsiung 807, Taiwan, R.O.C.
Tel: +886(0)7 3426031 Ext. 6327 E-mail: 97007@mail.wtuc.edu.tw

Abstract: Fuzzy time-series models have been utilized in making reasonably accurate predictions in many areas, such as academic enrollments, weather forecasting and stock markets. To refine past fuzzy time-series models, this paper proposes a new model, which employs the concepts of "momentum" along with Chebyshev's theorem in the forecasting process. The proposed model applies a "momentum" index to generate forecasting rules (fuzzy logical relationships) to reduce the probability of rules not being found in cases where no rules are available to forecast a testing dataset. Chebyshev's theorem is adopted to define a "reasonable" universe of discourse for the observations in a training dataset. From the refined process, two types of universe, symmetrical and asymmetrical, are given. To verify the proposed model, this paper employs experimental datasets, derived from a seven-year period of the Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX). Model comparison results show that the proposed model surpasses in accuracy one traditional fuzzy time-series model and two advanced models, based on neural networks and rough set algorithms.

JEL Classifications: C22, C43, C63, C13
Keywords: Fuzzy time-series, Stock price forecasting, Fuzzy linguistic variable

1. Introduction

With the emergence of data mining techniques, more and more artificial intelligence (AI) algorithms, such as fuzzy theory (Zadeh, 1975a;Zadeh, 1975b;Zadeh, 1976;Zimmermann, 1991), genetic algorithms (Bauer, 1994), and neural networks (Gately, 1996;Refenes et al., 1997;Trippi and Turban, 1993;Azo, 1994) have been applied to forecast complex systems, such as stock markets. In recent decades, many fuzzy time-series models, which utilize the fuzzy theory to model the relationships between the observations in time series, have been advanced to solve various forecasting problems, such as university enrollment forecasting (Song and Chissom, 1994;Song and Chissom, 1993b;Chen, 2002;Chen and Hsu, 2004;Chen and Chung, 2006;Chen, 1996;Song and Chissom, 1993a), weather prediction (Chen, 2000), and stock market forecasting (Huarng and Yu, 2006b;Huarng, 2001b;Huarng and Yu, 2005;Huarng and Yu, 2003;Yu, 2005a;Chen et al., 2008;Chen et al., 2007;Cheng et al., 2006;Su et al., 2010;Teoh et al., 2008;Cheng et al., 2010;Teoh et al., 2009;Huarng and Yu, 2006a;Yu and Huarng, 2008). For academic researchers, stock price forecasting is a very popular topic and, therefore, fuzzy time-series research usually employs stock market prices as prediction targets.

After reviewing the literature, two types of fuzzy time-series models can be summarized: (1) the traditional model, based on fuzzy theory (Huarng, 2001b;Huarng, 2001a;Huarng and Yu, 2005;Huarng and Yu, 2003;Teoh, et al., 2009;Yu, 2005a;Yu, 2005b;Chen et al., 2008;Chen et al., 2007;Chen, 2002;Chen and Hsu, 2004;Chen, 1996); and (2) the advanced model, based on fuzzy theory and other AI algorithms, such as genetic algorithms (Chen and Chung, 2006), neural

networks (Huarng and Yu, 2006b), and rough set algorithms (Teoh et al., 2009;Teoh et al., 2008;Su et al., 2010). Upon examining these models, we note that two issues, which may be critical points for forecasting performance, have rarely been discussed: (1) a "reasonable" universe of discourse for the observations in the training dataset; and (2) a probability reduction where "no rules"(fuzzy logical relationships) are available to forecast the testing dataset. To ascertain whether these problems effect forecasting performance for fuzzy time-series models, this paper proposes two approaches, "momentum" and Chebyshev's theorem, in order to refine the forecasting process.

To verify the proposed model, experimental datasets were selected from a 7-year period of the TAIEX. Further, four fuzzy time-series models, Chen's (1996), Huarng et al.'s (2005), Huarng et al.'s (2006) and Su et al.'s (2010) models, are used as comparison models to examine the forecasting improvement of the refined process. The remaining content of this paper is organized as follows: Section 2 introduces the definitions of a fuzzy time-series model; Section 3 offers the proposed concepts and algorithm; Section 4 submits model verification; Section 5 notes findings and suggests discussions; and the last section provides a conclusion.

2. Fuzzy Time-series Model

The fuzzy time-series, originally advanced by Song and Chissom (1993a; 1993b; 1994), is a data mining technique, which applies the fuzzy theory (Zadeh, 1975b;Zadeh, 1975a;Zadeh, 1976) to model fuzzy relationships among time-series observations. In the area of fuzzy time-series, numerous models have been proposed. Chen (1996) proposed another method to apply simplified arithmetic operations in forecasting algorithms rather than the complicated max-min composition operations presented in Song and Chissom's model. In his subsequent research, Chen proposed several methods (Chen, 2000;Chen, 2002;Chen and Hsu, 2004;Chen and Chung, 2006;Chen, 1996), such as high-order fuzzy relationships and genetic algorithms (Chen and Chung, 2006) to improve his initial model. Additionally, Huarng (2001a) pointed out that the length of intervals affected the forecasting accuracy in fuzzy time-series and proposed a method with a distribution-based length and an average-based length to reconcile this issue. In Huarng's model, two different lengths of intervals were applied to Chen's model, and it was concluded that the distribution-based and the average-based lengths could improve forecasting accuracy. Although this method demonstrates excellence in forecasting, we argue that it creates too many linguistic values to be identified by analysts, since, according to Miller(1956), establishing linguistic values and dividing intervals would be a trade off between human recognition and forecasting accuracy (Miller, 1956). Besides, Yu (2004) proposed a weighted model to tackle the two issues of recurrence and weighting in fuzzy time-series forecasting and concluded that the weighted model outperforms one of the conventional fuzzy time-series models (Yu, 2005a). The researcher argued that recurrent fuzzy relationships should be considered in forecasting, and he recommended that different weights be assigned to various fuzzy relationships.

In recent research, AI algorithms, such as genetic algorithms (Bauer, 1994), neural networks (Trippi and Turban, 1993) and rough set algorithms (Pawlak and Skoworn, 2007;Pawlak, 1982) were usually applied and efficiently improved the forecasting accuracy of fuzzy time-series models. Huarng and Yu (2006) applied neural networks to extract fuzzy logical relationships from fuzzy time-series to forecast the TAIEX (Huarng and Yu, 2006b) and their model outperformed Chen's (1996) model. Chen and Chung (2006) employed genetic algorithms to adjust the length of each interval in the universe of discourse (Chen and Chung, 2006). Teoh et al. (2009) adopted rough set algorithms to mine fuzzy logical relationships to enhance performance (Teoh et al., 2009;Teoh et al., 2008) and Su et al. (2010) used the algorithms to deal with technical indicators to produce forecasting rules for stock markets (Su et al., 2010). Although many types of fuzzy time-series models have been proposed to improve forecasting performance, the author believes that there are other reasonable approaches, such as stock analysis theory or techniques which can refine fuzzy time-series, to enhance forecasting accuracy whenever the forecasted targets are financial.

A fuzzy time-series can be defined as follows (Huarng and Yu, 2005;Song and Chissom, 1993a;Song and Chissom, 1994;Song and Chissom, 1993b;Yu, 2005b):

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

Definition 2: Suppose $F(t)$ is caused by $F(t\text{-}1)$ only. The relationship is expressed as $F(t) = 40\ F(t\text{-}1) * R(t,t\text{-}1)$; where $R(t,t\text{-}1)$ is the fuzzy relationship between $F(t)$ and $F(t\text{-}1)$; * represents an operator.

Let $F(t) = A_i$ and $F(t\text{-}1) = A_j$. The relationship between $F(t)$ and $F(t\text{-}1)$ (referred to as a fuzzy logical relationship, FLR(Song and Chissom, 1993a)) can be denoted by $A_i \rightarrow A_j$; where A_i is called the left-hand side (LHS) and A_j the right-hand side (RHS) of the FLR.

Definition 3: Given two FLR s with the same fuzzy sets on the LHSs, $A_i \rightarrow A_{j1}$; $A_i \rightarrow A_{j2}$. Both FLRs can be grouped together into fuzzy logical relationship groups (Chen, 1996): $A_i \rightarrow A_{j1}, A_{j2}$.

In general, a fuzzy time-series model includes the following forecasting processes (Yu, 2005b): (1) define the universe of discourse and the intervals for the observations; (2) partition the universe based on the intervals; (3) define the fuzzy sets for the observations; (4) fuzzify the observations; (5) establish the fuzzy relationship; (6) perform the forecast; and (7) defuzzify the forecasting results.

3. Momentum-based Fuzzy Time-series

After reviewing the literature of fuzzy time-series, there are two important issues, regarding the data spread, which are worth further discussion: (1) how to determine a reasonable universe of discourse for the stock price in the training datasets; and (2) how to provide an approach which avoids the condition of missing rules (fuzzy logical relationships) when the stock prices in a testing dataset are out of the defined boundary of its corresponding training dataset.

Firstly, the researcher defines the universe of discourse for the fuzzy time-series model training datasets subjectively (Yu, 2005a;Yu, 2005b;Chen et al., 2008;Chen et al., 2007;Chen, 2002;Chen and Hsu, 2004;Chen, 1996). These models set the boundaries for the universe of discourse with two unfixed and subjective numbers (see equation (1), where D_{min} is the minimum and D_{max} is the maximumin a training dataset; D_1 is a positive value given by the researcher) on the condition that the boundaries can cover the data distribution in the training sets. This is a less persuasive approach without any theory to support the process.

$$U = \left[D_{\min} - D_1, D_{\max} + D_1\right] \tag{1}$$

From Chebyshev's theorem (defined in equation (2), where x is a random variable; u is the mean of the random value x; k is a constant value; $\sigma$is the variance of the random value x) (Douglas et al., 2004), the probability of any random variable within a certain boundary can be determined. Therefore, to define the universe of discourse for training datasets with confirmable confidence, I argue that this theorem is a proper approach to refine the process of the traditional models in boundary defining.

$$P(u - k\sigma < x < u + k\sigma) \geq 1 - \frac{1}{k} \tag{2}$$

By applying the Chebyshev's theorem, two types of universes (U), asymmetrical and symmetrical, are provided to determine a reasonable universe of discourse for observations with a probability that is confirmable. The boundaries of the asymmetrical U are defined in equations (3) and (4). They are determined by two different adjustable parameters (k_1 and k_2), based on the maximum (D_{max}) and minimum (D_{min}) in a training dataset.

$$u - k_1 * s \leq D_{\min} \tag{3}$$

$$u + k_2 * s \geq D_{\max} \tag{4}$$

where u is the mean, s is the standard deviation, k_1 is an adjustable parameter to mark the low boundary (u-k_1* s) close to the minimum observation, k_2 is an adjustable parameter to mark the high boundary (u+k_2* s) close to the maximum observation.

However, the boundaries of the symmetrical U are determined by only one adjustable parameter (k_0), based on the maximum and minimum in the training dataset, and they are defined in equations (5) and (6).

$$u - k_0 * s \leq D_{\min} \tag{5}$$

$$u + k_0 * s \geq D_{\max} \tag{6}$$

where k_0 is an adjustable parameter, which marks the lower boundary (u-k_0* s), and the high boundary (u+k_0* s) close to the minimum and maximum.

Secondly, the experimental dataset used to verify fuzzy time-series is usually split into two sub-datasets, training and testing, with a specific ratio. In the forecasting process, problems sometimes occur when the observation in the testing dataset falls out of the boundaries defined by its corresponding training dataset. Whenever no rules are available for forecasting the testing dataset, a "naïve" forecast (denotes the use of the present observation as a forecast for the future) is employed to solve the problem (Chen, 2002;Chen and Hsu, 2004;Chen, 1996). However, when a stock market is in a full bear or full bull trend, there are no rules for forecasting the testing dataset and the naïve forecast completely replaces the predictions from the forecasting model, rendering the research model ineffective.

In the 2000 TAIEX, for example (see Figure 1), when the splitting ratio for the experimental dataset is set as 10:2 (a ten-month period of the stock index is selected for training, and the remainder for testing), many "naïve forecasts" are used because some observations in the testing dataset (ranged from 6090 to 4615) are much lower than the low boundary (5081) of the training dataset.

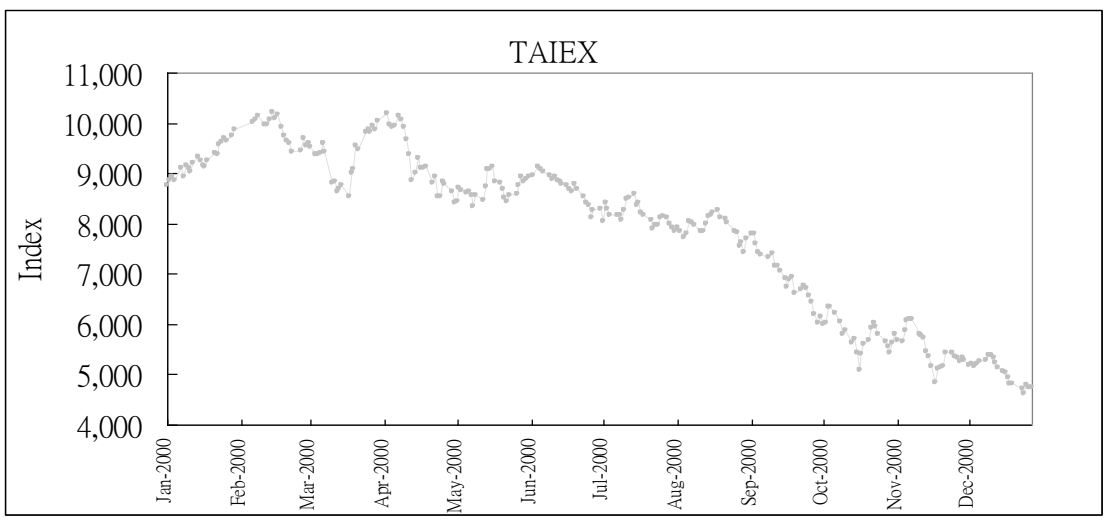

Figure 1 Index of the 2000 TAIEX

To overcome this problem, we introduce a momentum index. Stock price momentum is one of the technical indicators (Yamawaki and Tokuoka, 2007; Kim et al., 2006) used to measure price changes over a period of time. It is defined as equation (7), where m(t) is the stock price momentum, P(t), within a period of time, τ , which may vary by hours, days, months, etc., depending on the time length of the analysis (Wang H. and Pandey R.B., 2004).

$$m(t) = P(t+\tau) - P(t) \tag{7}$$

There are two reasons why this index is used: (1) daily price fluctuation limitations are sometimes imposed by governmental regulatory agencies, as the 7% limitation on the TAIEX; and (2) the volatility of the momentum is much smaller than the stock index in scale. Both can reduce the probability of the condition where no rules are available for forecasting a testing period. Taking a one-day momentum from the 2000 TAIEX as an example (see Figure 2), the observations in the testing dataset (from January to October) range from 274 to -323, and all of them are spread within the boundaries (-618,469) defined by the training dataset (from November to December).

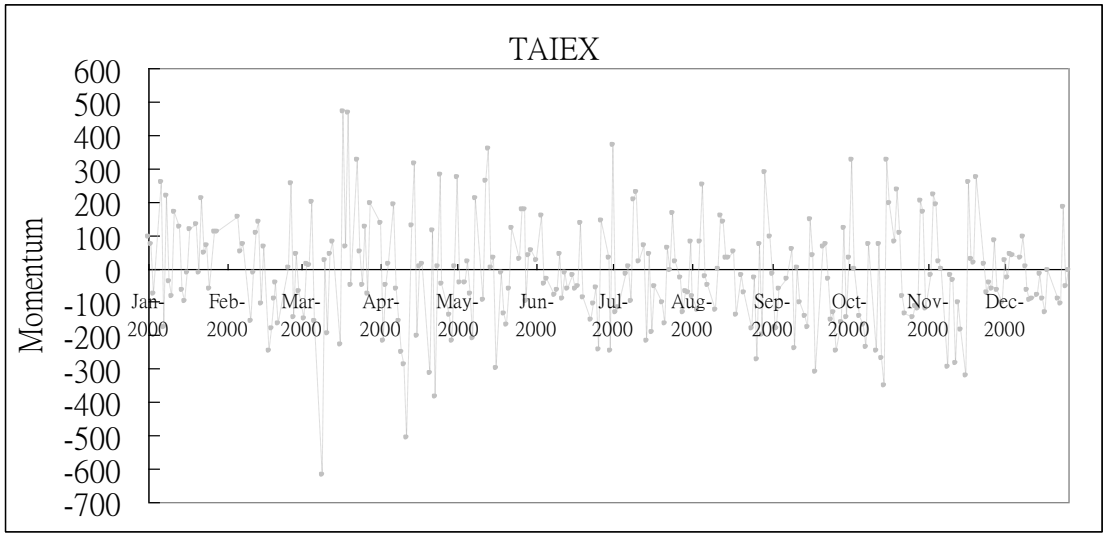

Figure 2 Momentums for the 2000 TAIEX

Based on these two refinements, a new fuzzy time-series model is proposed to improve forecasting performance. The proposed algorithm is introduced in the following section.

In this subsection, the proposed algorithm is introduced, step by step, using the 2000 TAIEX (see Figure 1 and Figure 2) as an example.

Step 1: Transfer stock index into momentum index

In the first step, the experimental dataset of the stock index is transferred into a dataset of a momentum index. The proposed algorithm refines equation (7) as equation (8) to produce a momentum index as a forecasting factor because equation (7) includes the unknown information, P(t+τ), which denotes the future stock index within a period of time τ (τ is set as "1" in this paper).

$$m(t) = P(t) - P(t-\tau) \tag{8}$$

Step 2: Reasonably define the universe of discourse for the observations

In this step, equations (3) to (5) are employed to define two types of the universe of discourse: asymmetrical U and symmetrical U. The high and low boundaries of symmetrical U and asymmetrical U are given by equation (9) and equation (10), respectively.

$$\text{symmetrical U} = [u - k_0 * s, u + k_0 * s] \tag{9}$$

$$\text{asymmetrical U} = [u - k_1 * s, u + k_2 * s] \tag{10}$$

For example, the high and low boundaries of asymmetrical U and symmetrical U for the training dataset, employing the 2000 TAIEX (from January to October), are listed in Table 1. The mean (u) for the observations is -14.41, standard deviation(s) is 163.19, the minimum and maximum is -618 and 469, k_0 is 3.7, k_1 is 3.7, and k_2 is 3.

Table 1 The high and low boundaries of asymmetrical *U* and symmetrical *U*

Defined *U*		Year (Mean)	2000 (-14.41)
Symmetrical *U*	$[u - k_0*s, u + k_0*s]$		[-618,589]
Asymmetrical *U*	$[u - k_1*s, u + k_2*s]$		[-618,475]

Step 3: Partition the universe, based on the intervals

This step partitions the defined universe from Step 2 into a certain number of intervals for observations in the training datasets. The proposed algorithm uses seven intervals to partition the defined universe, as listed in Table 2.

Table 2 Seven partitioning intervals for observations

Symmetrical *U*		**Asymmetrical *U***	
Intervals	Range	Intervals	Range
I_1	[-618, -446]	I_1	[-618,-462]
I_2	[-446, -273]	I_2	[-462,-306]
I_3	[-273, -101]	I_3	[-306, -150]
I_4	[-101, 72]	I_4	[-150, 7]
I_5	[72, 244]	I_5	[7, 163]
I_6	[244, 417]	I_6	[163, 319]
I_7	[417, 589]	I_7	[319, 475]

Step 4: Define the fuzzy sets for the observations

This step establishes a related fuzzy set (linguistic value) for each observation in the training dataset. The fuzzy sets, $A_1\ A_2 \ldots\ A_k$, for the universe of discourse are defined by equation (11).

$$\begin{aligned} A_1 &= a_{11}/u_1 + a_{12}/u_2 + \ldots + a_{1m}/u_m \\ A_2 &= a_{21}/u_1 + a_{22}/u_2 + \ldots + a_{2m}/u_m \\ &\vdots \\ A_k &= a_{k1}/u_1 + a_{k2}/u_2 + \ldots + a_{km}/u_m \end{aligned} \tag{11}$$

For example, in this paper, the seven fuzzy linguistic values A_1 = (very low momentum), A_2 = (low momentum), A_3 = (slightly low momentum), A_4 = (zero momentum), A_5 = (slightly high momentum), A_6 = (high momentum) and A_7 = (very high momentum) are refined by equation (12) (Chen, 1996).

$$\begin{aligned} A_1 &= 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + 0/u_5 + 0/u_6 + 0/u_7 \\ A_2 &= 0.5/u_1 + 1/u_2 + 0.5/u_3 + 0/u_4 + 0/u_5 + 0/u_6 + 0/u_7 \\ A_3 &= 0/u_1 + 0.5/u_2 + 1/u_3 + 0.5/u_4 + 0/u_5 + 0/u_6 + 0/u_7 \\ A_4 &= 0/u_1 + 0/u_2 + 0.5/u_3 + 1/u_4 + 0.5/u_5 + 0/u_6 + 0/u_7 \\ A_5 &= 0/u_1 + 0/u_2 + 0/u_3 + 0.5/u_4 + 1/u_5 + 0.5/u_6 + 0/u_7 \\ A_6 &= 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0.5/u_5 + 1/u_6 + 0.5/u_7 \\ A_7 &= 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0/u_5 + 0.5/u_6 + 1/u_7 \end{aligned} \tag{12}$$

Step 5: Fuzzify the observations

In equation (12), the value of a_{ij} indicates the grade of membership of u_j in fuzzy set A_i, where $a_{ij} \in [0,1]$, $1 \leq i \leq k$ and $1 \leq j \leq m$; and ascertains the degree of each stock price belonging to each A_i ($i=1,\ldots,m$). If the maximum membership of the stock price is under A_k, then the fuzzified stock price is labeled as A_k. Each observation in the training dataset can be classified by the seven partitioned intervals and labeled with a specific linguistic value based on equation (11). Table 3 demonstrates the assignments of linguistic value with symmetrical U for eight periods of momentum indices, based on the classifications from equation (12).

Table 3 Assign a related linguistic value to each momentum index (Symmetrical U)

Date	Time sequence	Stock index	Momentum index	Linguistic value
01/04/2000	$t = 0$	8,756.55		
01/05/2000	$t = 1$	8,849.87	93.32	A_5
01/06/2000	$t = 2$	8,922.03	72.16	A_5
01/07/2000	$t = 3$	8,845.47	-76.56	A_4
01/10/2000	$t = 4$	9,102.60	257.13	A_6
01/11/2000	$t = 5$	8,927.03	-175.57	A_3
01/12/2000	$t = 6$	9,144.65	217.62	A_5
01/13/2000	$t = 7$	9,107.19	-37.46	A_4
01/14/2000	$t = 8$	9,023.24	-83.95	A_4

Step 6: Establish the fuzzy relationship

In this step, there are three procedures: (1) establish a fuzzy logic relationship (FLR) for linguistic time-series; (2) summarize all fuzzy logic relationships as FLR groups to produce a rule matrix; and (3) assign a weight to each FLR group. Firstly, one FLR (Fuzzy Logical Relationship) is composed of two consecutive linguistic values. For example, the FLR $A_i \rightarrow A_j$ is established by A_i (t-1) and A_j (t). Table 4 demonstrates the FLR, establishing processes of the linguistic time-series from Table 3.

Secondly, the FLRs with the same LHS (Left Hand Side) linguistic value can be grouped into one FLR group such as $A_i \rightarrow A_{j1}, A_{j2}$. All FLR groups will construct a rule matrix. Table 5 demonstrates the matrix produced by the FLRs from Table 4. Each row of the matrix represents one FLR group and each cell represents the occurrence frequency of each FLR.

Thirdly, each FLR within the same FLR group should be assigned a weight. In the proposed algorithm, the trend-weighted method is applied (Cheng, Chen, and Chiang, 2006). For example, in Table 5, the FLR group of A_5 is $A_5 \rightarrow A_4, A_5$. The FLR of $A_5 \rightarrow A_5$ occurs once and the weight is assigned 1. However, The FLR of $A_5 \rightarrow A_4$ occurs twice. Therefore, the first FLR is assigned 1 and the second FLR is assigned 2. In this method, the FLR weight is determined by the order of occurrence. The sum of the weight of each FLR should be standardized to obtain a trend-weighted rule matrix, defined as W_n(t) (see Equation (13)).

Table 4 FLR Table

$A_5(t = 1) \rightarrow A_5(t = 2)$
$A_5(t = 2) \rightarrow A_4(t = 3)$
$A_4(t = 3) \rightarrow A_6(t = 4)$
$A_6(t = 4) \rightarrow A_3(t = 5)$
$A_3(t = 5) \rightarrow A_5(t = 6)$
$A_5(t = 6) \rightarrow A_4(t = 7)$
$A_4(t = 7) \rightarrow A_4(t = 8)$

Table 5 Example of a rule matrix from FLRs

$m(t)$ \ $m(t+1)$	A_1	A_2	A_3	A_4	A_5	A_6	A_7
A_1	0	0	0	0	0	0	0
A_2	0	0	0	0	0	0	0
A_3	0	0	0	0	1	0	0
A_4	0	0	0	1	0	1	0
A_5	0	0	0	2	1	0	0
A_6	0	0	1	0	0	0	0
A_7	0	0	0	0	0	0	0

$$W_n(t) = [W_1', W_2', \ldots, W'_i] = \left[\frac{W_1}{\sum_{k=1}^{i} W_k}, \frac{W_2}{\sum_{k=1}^{i} W_k}, \cdots, \frac{W_i}{\sum_{k=1}^{i} W_k} \right] \quad (13)$$

For example, the standardized weights for A_5 in Table 5 are specified as follows: W_1 =0, W_2 = 0, W_3 = 0, W_4 = 3/4, W_5 = 1/4, W_6 = 0 and W_7 = 0. For example, Table 6 demonstrates the trend-weighted rule matrix converted from Table 5 above.

Table 6 Example of trend-weighted rule matrix

$m(t)$ \ $m(t+1)$	A_1	A_2	A_3	A_4	A_5	A_6	A_7	$\sum_{k=1}^{i} W_k$
A_1	0	0	0	0	0	0	0	0
A_2	0	0	0	0	0	0	0	0
A_3	0	0	0	0	1	0	0	1
A_4	0	0	0	1/2	0	1/2	0	2
A_5	0	0	0	3/4	1/4	0	0	4
A_6	0	0	1	0	0	0	0	1
A_7	0	0	0	0	0	0	0	0

Step 7: Perform the forecast

With the rule matrix from Step 6, the forecasts for the testing dataset can be generated by the following statement: suppose m(t) = A_i . The forecast of $m(t+1)$ is produced by the following rules:

(1) *if* A_i is not found in the FLR groups (the rule matrix from Step 6), the forecast of $m(t+1)$ is equal to m(t).

(2) *if* A_i = the LHS (Left Hand Side) of a FLR group, such as $A_i \rightarrow A_{j1}, A_{j2}, \ldots, A_{jk}$, the forecast of $m(t+1)$ is $W_1 * A_{j1}, + W_2 * A_{j2}, \ldots, + W_k * A_{jk}$.

Step 8: Defuzzify the forecasting results

In this step, the linguistic forecast from Step 7 is converted to a numerical forecast, which is called "defuzzification" as in the following statement: *if* the forecast of $m(t+1)$ is $W_1 * A_{j1}, + W_2 * A_{j2}, \ldots, + W_k * A_{jk}$, the defuzzified forecasting result is equal to a trend-weighted arithmetic average of the midpoints for $A_{j1}, A_{j2}, \ldots, A_j$ (Chen, 1996) and defined by equation (14).

$$forecast\_m(t+1) = \mathrm{W_1 * MP(A_{j1}) + W_2 * MP(A_{j2}), \ldots, + W_k * MP(A_{jk})} \quad (14)$$

Step 9: Transfer the momentum index to the stock index

The defuzzified forecasting result from Step 8 is a momentum index and has to be converted to a stock index. This step employs equation (15) to produce a forecast of the stock index for the next day.

$$forecast\_P(t+1) = P(t) + forecast\_m(t+1) \quad (15)$$

4. Model Verification

In this section, a seven-year period of the TAIEX, from 1999/1/4 to 2005/12/31, is used to create experimental datasets (the first ten-month period of each year, from January through October, is used for training and the last two, November and December, for testing), and to evaluate forecasting performance, the RMSE (Root Mean Square Error, defined in equation (16) on the next page) is employed as a performance indicator.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} \left| actual(t) - forecast(t) \right|^2}{n}} \tag{16}$$

4.1 Performance Evaluation

Since the data spread is a key issue for the proposed model, the statistics of the observations are the referenced factors to define the universe of discourse, *U*, for the stock price in the training datasets. Table 7 lists the statistics for the training periods from 1999 to 2003. By applying Chebyshev's theorem, the asymmetrically and symmetrically defined universes, which evaluate the forecasting performance with confirmable confidence of probability, are provided. Table 7 shows the forecasting performance for testing periods with two types of defined universes. It is clear that the types of boundaries vary the performance under the same forecasting processes.

4.2 Model Comparison

To verify the forecasting performance of the proposed model, several time-series models with one forecasting factor (stock index) are employed as comparison models. Using the performance data from Huarng et al.'s (2005; 2006b), Yu et al.'s (2008) and Su et al.'s (2010) research, a performance comparison table is produced and listed in Table 9. It is apparent that the proposed model with two types of defined U bears the smallest RMSE for every testing period.

Table 7 Statistics of momentum for training periods

Statistics \ Year	1999	2000	2001	2002	2003	2004	2005
Mean	7.74	-14.41	-5.16	-5.00	7.42	-1.65	-1.88
Standard deviation	118	163	91	96	69	96	48
Variance	13839	26632	8342	9141	4759	9220	2336
Kurtosis	2.29	0.83	0.05	0.29	0.43	3.76	0.89
Skewness	-0.07	0.04	0.31	0.14	0.05	-0.57	-0.16
Range	926.42	1086.08	507.72	559.47	401.08	796.66	304.54
Minimum	-506.46	-617.65	-224.44	-284.22	-189.99	-455.17	-173.21
Maximum	419.96	468.43	283.28	275.25	211.09	341.49	131.33
* Difference of spread distance from Mean	101.98	120.4	69.16	1.03	6.26	110.38	38.12

Note: "*" is defined by the equation: $\left\| \min imum - mean \right| - \left| \max imum - mean \right\|$

Table 8 Forecasting performance for testing periods with two types of defined *U*

Defined *U* \ Year (Mean)		1999 (7.74)	2000 (-14.41)	2001 (-5.16)	2002 (-5.00)	2003 (7.42)	2004 (-1.65)	2005 (-1.88)
Sym. *U*	$[u + k_0 * s, u - k_0 * s]$	[-508,523]	[-618,589]	[-297,287]	[-286,276]	[-197,212]	[-456,453]	[-173,170]
	RMSE	*103	130	*120	*68	*55	56*	54
Asym. *U*	$[u + k_1 * s, u - k_2 * s]$	[-510,431]	[-618,475]	[-233,287]	[-285,276]	[-191,212]	[-456,342]	[-173,132]
	RMSE	109	*122	125	*68	58	58	53*

Note: * indicates smaller RMSE.

Table 9 Probability that observation falling into defined U based on Chebyshev's theorem

Defined *U* \ Year (Mean)		1999 (7.74)	2000 (-14.41)	2001 (-5.16)	2002 (-5.00)	2003 (7.42)	2004 (-1.65)	2005 (-1.88)
Sym. *U*	[k_0]	[4.38]	[3.7]	[3.2]	[2.94]	[2.96]	[4.73]	[3.55]
	Prob.	94.79%	92.70%	90.23%	88.43%	88.59%	95.53%	92.07%
Asym. *U*	[k_1, k_2]	[4.38,3.51]	[3.7,3.0]	[2.5,3.2]	[2.93,2.94]	[2.87,2.96]	[4.73,3.58]	[3.55,2.76]
	Prob.	93.34%	90.79%	87.12%	88.39%	88.22%	93.86%	89.47%

Table 10 Performance comparisons for testing periods (TAIEX)

Models \ **Year**		1999	2000	2001	2002	2003	2004	2005
**Conventional regression model		164	420	1070	116	329	146	--
Chen's model (1996)		120	176	148	101	74	83	66
Huarng et al.'s model (2005)		--	139	144	82	73	--	--
Huarng et al.'s model (2006b)		109	152	130	84	56	79	69
Su et al.'s model (2010)		--	--	122	94	55	69	65
Proposed	(Sym. *U*)	*103	130	*120	*68	*55	*56	54
	(Asym. *U*)	109	*122	125	*68	58	58	*53

Note: "*" denotes minimum RMSE; "--"denotes no datum;

"**" the performance data are referred from Yu et al. (2008).

4.3 Findings and Discussions

After implementing the experiment, several findings were discovered from the experimental data which are summarized in Tables 7 through Table 10 and introduced below.

Firstly, the larger the variance of the stock market leads, the poorer the forecasting performance. Table 8 shows that the RMSE of the proposed model from 1999 to 2001 (*Sym. U*: 103,130,120; *Asym. U*: 109,122,125) is worse than that for the period from 2002 to 2005 (*Sym. U*: 68,55,56,54; *Asym. U*: 68,58,58,53). From Table 7, the standard deviation of the observations (momentum index) from 1999 to 2001(118,163,91) is almost larger than that for the period from 2002 to 2005 (96,69,96,48). Although there is one exception in 2001, the relation between the variance and the forecasting performance is still obvious.

Secondly, the spread characteristics of observations influence the forecasting performance of fuzzy time-series. From Table 8, it is apparent that the RMSE for the two types of defined *U* (Sym. *U* and Asym. *U*) is different, except for the year 2002 (where *Sym. U* is close to *Asym. U*). Because the two types of defined *U* are given according to the spread characteristics of observations, it is believed that forecasting performance of fuzzy time-series is related to observation spread. Besides, the RMSE performs better employing an Asym. *U* when the observations spread into extreme

asymmetry. The phenomenon is shown in the TAIEX for year 2000 (the difference of the spread distance from Mean is 120, which is the largest among the 7 experimental datasets, and the RMSE for an Asym. *U* is much better than for a Sym. *U*).

Thirdly, the probability that the observations in the testing dataset fall into the two types of defined *U* can boost the confidence level in fuzzy time-series models, defining the universe for observations. Based on Chebyshev's theorem, the least probable of the two types of U can be produced and listed in Table 9. This can provide useful information for researchers to decide whether the universe for the fuzzy time-series model should be modified or not (increased or decreased) to reach its maximum forecasting performance.

Lastly, the momentum index is a better forecasting factor than the stock index for fuzzy time-series. The comparison results of Table 10 show that the proposed model, based on a momentum index, performs better than the four fuzzy time-series models (Chen, S.M., 1996; Huarng et al., 2005; Huarng et al., 2006b; Su et al., 2010) and a regression model, all of which are based on a daily stock index. The reason that the momentum index is a better forecasting factor can be explained by its ability to reduce the variability of a forecasting factor for the stock market, and provide more meaningful information for stock investor than daily stock indices.

5. Conclusions and Further Works

In this paper, a new fuzzy time-series, which employs a momentum index and Chebyshev's theorem to define a universe, has been proposed in order to improve forecasting accuracy. From the verification results, it is concluded that the major research goal has been reached. Additionally, two important findings have been discovered: (1) a larger variance of the stock market leads to poorer forecasting performance; and (2) the spread characteristics of observations influence the forecasting performance of fuzzy time-series. Besides the conclusion, the proposed model can provide a practical application in investment decisions. Investors can make their decisions, buying or selling stocks, based on the forecasting momentum index for the next day. For example, if the forecasting index is positive, investors should buy the stock index at the opening price and sell it at the closing price to make a profit; conversely, they should sell the index at the opening price and buy it back when the forecast is negative.

However, there is still room for testing and improving this model, as follows, e.g.: (1) employing other stock markets as testing datasets to evaluate the performance; (2) simulating the model to trade in the stock market, and sum up the profits of these trades to evaluate profit making; and (3) reconsidering the factors affecting the behaviour of the stock markets, such as trading volume, news and financial reports.

Acknowledgements: This work was supported in part by the National Science Council, Republic of China (Taiwan), under Grant NSC 100-2221-E-160-001.

References

[1] Azo, M. E. (1994), *Neural Network Time Series Forecasting of Financial Markets*, New York: Wiley.

[2] Bauer Jr., R. J. (1994), *Genetic Algorithms and Investment Strategies*. New York: Wiley.

[3] Bollerslev T. (1986), "Generalized autoregressive conditional heteroscedasticity", *Journal of Econometrics*, 31(3): 307-327.

[4] Box, G. and Jenkins, G.(1976), *Time series analysis: Forecasting and control*, San Francisco: Holden-Day.

[5] Chen, T.L, Cheng, C.H., and Teoh H.J. (2008), "High Order Fuzzy Time-series Based on Multi-period Adaptation Model for Forecasting Stock Markets", *Physica A*, 387(4): 876-888.

[6] Chen T.L., Cheng, C.H., and Teoh H.J.(2007), "Fuzzy Time-series Based on Fibonacci Sequence for Stock Price Forecasting", *Physica A*, 380(2): 377-390.

[7] Chen S. M. (1996), "Forecasting enrollments based on fuzzy time-series", *Fuzzy Sets and Systems*, 81(3): 311-319.

[8] Chen S. M. (2000), "Temperature Prediction Using Fuzzy Time-series", *IEEE Trans.CYBERNETICS*, 30(2): 263-275.

[9] Chen S. M. (2002), "Forecasting enrollments based on high-order fuzzy time series", *Cybernetics and Systems*, 33(1): 1-16.

[10] Chen, S. M. and Chung N. Y. (2006), "Forecasting enrollments using high-order fuzzy time series and genetic algorithms", *International Journal of Intelligent Systems*, 21(5): 485-501.

[11] Chen, S. M. and Hsu C. C. (2004), "A new method to forecast enrollments using fuzzy time series", *International Journal of Applied Science and Eng.*, 5(2,3): 234-244.

[12] Cheng, C. H., Chen, T. L., and Chiang C. H. (2006), "Trend-Weighted Fuzzy Time-Series Model for TAIEX Forecasting", *Lecture Notes in Computer Science*, Vol.4234: 469-477.

[13] Cheng, C. H., Chen, T. L., and Wei L. Y. (2010), "A Hybrid Model Based on Rough Sets Theory and Genetic Algorithms for Stock Price Forecasting", *Information Sciences*, 180(9): 1610-1629.

[14] Douglas C.M., George C.R., and Norma, F.H. (2004), *Engineering statistics*, New York: Wiley.

[15] Gately, E.(1996), *Neural Networks for Financial Forecasting*, New York: John Wiley and Sons.

[16] Hanke, J.E. and Reitsch, A.G. (1995), *Business forecasting*. New Jersey: Prentice-Hall.

[17] Huarng K. H. (2001a), "Effective lengths of intervals to improve forecasting in fuzzy time series", *Fuzzy Sets and Systems*, 123(3): 387-394.

[18] Huarng K. H. (2001b), "Heuristic models of fuzzy time series for forecasting", *Fuzzy Sets and System*, 123(3): 369-386.

[19] Huarng, K. H. and Yu H. K. (2003), "An N-th order heuristic fuzzy time series model for TAIEX forecasting", *International Journal of Fuzzy Systems*, 5(4): 247-253.

[20] Huarng, K. H. and Yu H. K. (2005), "A type-2 fuzzy time-series model for stock index forecasting", *Physica A*, Vol.353: 445-462.

[21] Huarng, K. H. and Yu T. H. K.(2006a), "Ratio-Based Lengths of Intervals to Improve Fuzzy Time Series Forecasting", *IEEE Transactions On Systems, Man, And Cybernetics-Part B: Cybernetics*, 36(2): 328-340.

[22] Huarng, K. H. and Yu T. H. K. (2006b), "The application of neural networks to forecast fuzzy time series", *Physica A*, 363(2): 481-491.

[23] Kim M.J., Min S.H., and Han I. (2006), "An evolutionary approach to the combination of multiple classifiers to predict a stock price index", *Expert Systems With Applications*, 32(2): 241-247.

[24] Miller G.A. (1956), "The magical number seven, plus or minus two: Some limits on our capacity of processing information", *The Psychological Review*, 63(2): 81-97.

[25] Pawlak Z. (1982), "Rough sets", *International Journal Computer and Information Science*, 11(5): 341-356.

[26] Pawlak, Z. and Skoworn A. (2007), "Rudiments of rough sets", *Information Sciences*, 177(1): 3-27.

[27] Refenes, A. N., Burgess, N., and Bentz Y. (1997), "Neural networks in financial engineering: a study in methodology", *IEEE Trans.Neural Networks*, 8(6): 1222-1267.

[28] Song, Q. and Chissom B. S. (1993a), "Fuzzy time-series and its models", *Fuzzy Sets and Systems*, 54(3): 269-277.

[29] Song, Q. and Chissom B. S. (1993b), "Forecasting enrollments with fuzzy time-series---Part Ⅰ", *Fuzzy Sets and Systems*, 54(1): 1-10.

[30] Song, Q. and Chissom B. S. (1994), Forecasting enrollments with fuzzy time-series Part Ⅱ. *Fuzzy Sets and Systems*, 62(1): 1-8.

[31] Su C.H., Chen T.L., Cheng C.H., and Chen Y.C. (2010), "Forecasting the Stock Market with Linguistic Rules Generated from the Minimize Entropy Principle and the Cumulativefrom the Minimize Entropy Principle and the Cumulative Probability Distribution Approaches", *Entropy*, 12(12): 2397-2417.

[32] Teoh H.J., Chen T.L., Cheng C.H., and Chu H.H. (2009), "A hybrid multi-order fuzzy time series for forecasting stock market", *Expert Systems With Applications*, 36(4): 7888-7897.

[33] Teoh, H. J., Cheng, C. H., Chu, H. H., and Chen, J. S. (2008), "Fuzzy time series model based on probabilistic approach and rough set rule induction for empirical research in stock markets", *Data and Knowledge Engineering*, 67(1): 103-117.

[34] Trippi, R. R. and Turban, E. (1993), *Neural Network in Finance and Investing*, New York: Probus Publishing Company.

[35] Wang H. and Pandey R.B. (2004), "A momentum trading approach to technical analysis of Dow Jones industrials", Ph*ysica A*, 331(3,4): 639-650.

[36] Yamawaki M.T. and Tokuoka S. (2007), "Adaptive use of technical indicators for the prediction of intra-day stock prices", Ph*ysica A: Statistical Mechanics and Its Applications*, 383(1): 125-133.

[37] Yu H. K. (2005), "Weighted fuzzy time-series models for TAIEX forecasting", Ph*ysica A*, 349(3,4): 609-624.

[38] Yu T. H. K. (2005), "A refined fuzzy time-series model for forecasting", Ph*ysica A*, 346(3, 4): 657-681.

[39] Yu, T. H. K. and Huarng K. H. (2008), "A Bivariate Fuzzy Time Series Model to Forecast the TAIEX", E*xpert Systems With Applications*, 34(4): 2945-2952.

[40] Zadeh L. A. (1975a), "The concept of a linguistic variable and its application to approximate reasoning -- Ⅰ", *Information Sciences*, 8(3): 199-249.

[41] Zadeh L. A. (1975b), "The concept of a linguistic variable and its application to approximate reasoning-- Ⅱ", *Information Sciences*, 8(4): 301-357.

[42] Zadeh L. A. (1976), "The concept of a linguistic variable and its application to approximate reasoning-- Ⅲ", *Information Sciences*, 9(1): 43-80.

[43] Zimmermann, H. J. (1991), *Fuzzy set theory and its applications*. Boston: Kluwer Academic Publishers.