Information Sciences 294 (2015) 227–241

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

A hybrid fuzzy time series model based on granular computing for stock price forecasting

Mu-Yen Chen *, Bo-Tsuen Chen

Department of Information Management, National Taichung University of Science and Technology, Taiwan

ARTICLE INFO

Article history:
Received 9 December 2012
Received in revised form 13 September 2014
Accepted 24 September 2014
Available online 13 October 2014

Keywords:
Granular computing
Binning-based partition
Entropy-based discretization
Time series forecasting

ABSTRACT

Given the high potential benefits and impacts of accurate stock market predictions, considerable research attention has been devoted to time series forecasting for stock markets. Over long periods, the accuracy of fuzzy time series model forecasting is invariably affected by interval length, and formulating effective interval partitioning methods can be very difficult. Previous studies largely relied on distance partitioning, but this approach neglects the distribution of datasets and can only handle scalar forecasting. But the magnitude of stock price movements is often severe and difficult to predict. Thus, the distribution of stock price datasets is always skewed and the straightforward partitioning method is not well suited to these types of time series datasets. In this research, a novel fuzzy time series model is used to forecast stock market prices. The proposed model is based on the granular computing approach with binning-based partition and entropy-based discretization methods. The proposed model is verified using experimental datasets from the Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX), Dow-Jones Industrial Average (DJIA), S&P 500 and IBOVESPA stock indexes, and results are compared against existing fuzzy time series models, three different SVM models, and three modern economic models - GARCH, GJR-GARCH, and Fuzzy GARCH. Compared to other current forecasting methods, the proposed models provide improved prediction accuracy and the results are verified by paired two-tailed *t*-tests. The experimental results clearly provide improvements for obtaining optimized linguistic intervals and ensuring the accuracy of the proposed model.

1. Introduction

The ongoing global financial crisis underlines the importance of the time series concept, and methods are urgently needed to analyze time series data from daily closing prices with the aim of accurately predicting stock price movements. Currently, financial forecasting relies mainly on mathematical and statistical methods [1,2,20,43,57], and time series models [3,4]. The 2007 financial crisis increased the importance of financial forecasting among general investors and researchers, prompting the development and application of many theories and techniques in fundamental and technical analysis.

A time series is defined as a set of sequential observations which can be either *continuous* or *discrete*. Time series analysis is widely used by researchers studying methods of stock market forecasting (originally proposed by Kendall and Ord [34]) and logistical regression models based on traditional statistical assumptions. Financial forecasting problems are usually

* Corresponding author at: 129 Sec. 3, San-Min Road, Taichung 404, Taiwan. Tel.: +886 4 22196310; fax: +886 4 22196311.
E-mail address: mychen@nutc.edu.tw (M.-Y. Chen).

http://dx.doi.org/10.1016/j.ins.2014.09.038
0020-0255/

handled using traditional time series methods such as autoregressive moving average (ARMA) model [4], autoregressive conditional heteroscedastic (ARCH) model [19], and Generalized ARCH (GARCH) model [3], but these methods require extensive historical data and assumptions such as normality postulates [33]. Unlike conventional time series which deal with real numbers, fuzzy time series are architecture by fuzzy sets [62]. Fuzzy time series are frequently applied into producing stock price predictions due to the handling capability of linguistic value datasets to produce accurate forecasting results. Nowadays, it has been widely and successfully utilized to forecast nonlinear and dynamics datasets in such widely varying domains including course enrollment [9,25,50,52], temperature [56], traffic accidents [33], tourism demand [55] and stock markets [13,14,28,59,60].

First proposed by Zadeh (1965), fuzzy set theory is still widely used in a broad range of applications. Song and Chissom [50] developed a fuzzy time series model to predict university enrollment levels. However, this model used max–min composition operations which significantly complicate the computational process. Chen [7] presented one simplified calculation process to solve this drawback, but this model lacked a suitable weight mechanism for the fuzzy logical relationships (FLRs). The model has since been widely extended and adapted, with researchers seeking to improve forecasting accuracy by adjusting the length of linguistic intervals or by changing the weighting approach. Huarng (2001) determined interval lengths using distribution-based and average-based lengths [25]. Yu (2005) applied different weighted technologies on fuzzy time series models to strength forecasting accuracy [59]. Huarng and Yu (2006) added the ratio-based interval length to the fuzzy time series model [24] while Cheng et al. (2006) integrated a fuzzy time series model with a trend-weighting mechanism to predict actual trading data of stock prices and university enrollment [12]. They later predicted hospital outpatient traffic adopting a revised fuzzy time series which integrated a weighted-transitional matrix [15]. These methods provide lower forecasting error rates did the methods proposed by Chen and Chung [9] and Yu [59].

These aforementioned studies all focus on linguistic interval length for fuzzy time-series models, which raises several problems: (1) while distance partitioning is the most widely used method for interval partitioning [9,37,45,49–52,65], it is unable to accurately reflect the distribution of authentic data and (2) the difficulty of defining and selecting reliable interval lengths leaves existing methods unable to provide adequate prediction accuracy [10,13]. These two drawbacks created problems for fuzzy time series forecasting models. First, distance partitioning sets each linguistic interval range to a uniform width, but this can easily result in either excessive linguistic values [60] or excessively short intervals which can lead to the generation of null sets among the FLRs. In fact, the magnitude of stock price movements is often severe and difficult to predict. Thus, the distribution of stock price datasets is always skewed, and is thus better suited to the frequency partitioning method. Secondly, reliable interval lengths are hard to obtain from dynamic and complex time series models (e.g., daily stock prices) [10,13]. However, the entropy-based discretization method has better results with datasets with continuous attributes than unsupervised discretization methods [17,18].

Granular computing (GrC) techniques can provide appropriate solutions to the limitations of fuzzy time series. GrC applies concepts including value discretization, group aggregation, classification or universe clustering to solve a range of problems [63]. To sum up, GrC is suddenly becoming an emerging technology and can be viewed as a superset of fuzzy information granulation, rough sets and interval computations [46]. From this perspective, the main research issues of granular computing for data mining can be categorized as belonging to three types: rule representation, rule mining, and soft computing integration [58]. Rule representation means one fuzzy granule can be determined to generalize constraints and can be represented in natural language. Rule mining refers to the capacity of GrC to obtain more general rules by grouping attributes into granules. It can also be used to investigate semantic relationships among these attributes through establishing a hierarchy of granules. Finally, soft computing integration refers to combining GrC with evolutionary computation or bio-inspired computation to significantly improve forecasting performance. Here, this research proposed one novel fuzzy time series method to forecast the stock market prices through binning-based partitioning and entropy-based discretization methods to obtain reliable interval lengths from skewed datasets. In addition, the entropy-based discretization method can also adjust the length of the individual intervals for each iteration and obtain reliable interval lengths. The proposed model was applied to experimental datasets taken from the Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX), Dow-Jones Industrial Average (DJIA), S&P 500, and IBOVESPA stock indexes. Results were then compared against the models suggested by Chen (1996) [7], Yu (2005) [59], Chang et al. (2011) [5], Hsieh et al. (2011) [23], Chen and Chen (2011) [8], Chen et al. (2011) [11], Cheng et al. (2013) [16], Chen and Kao (2013) [10], and three support vector regression (SVR) models with different kernel functions [53,54]. This research also compared our model to current economic models including GARCH [3], GJR-GARCH (Glosten–Jagannathan–Runkle GARCH) [21], and Fuzzy GJR-GARCH [31].

The structure of this research is illustrated as follows. Section 2 investigates relevant studies regarding fuzzy time series definitions, fuzzy time series models, and entropy concepts. The research methodology is described in Section 3. In Section 4, the experimental design and experimental results are discussed. Finally, Section 5 summarizes the conclusions and potential issues in the future researches.

2. Literature review

Solutions to strategic decision problems rely heavily on accurate forecasting to reduce losses from poor decisions and to increase firm competitiveness. Modern firms are increasingly reliant on fuzzy time series for forecasting, and this research reviews theories related to fuzzy time series and related interval computing methods.

2.1. Fuzzy time series models

The last decade has seen a significant expansion on the original fuzzy time series concept. These modifications can be categorized into high-order models [11,16,38,39,45,48] and soft computation approaches [5,10,14,23,37,61]. The relative simplicity of first-order fuzzy time series models frequently leaves them unable to explain more complicated relationships [26,27]. Own and Yu (2005) proposed a heuristic, high-order fuzzy time series model to deal with prediction issues [45]. Li and Cheng (2007) presented a deterministic model to produce university enrollment forecasts which outperformed existing conventional models [38]. Singh (2009) used various sequences as forecasting parameters and employed a w-step fuzzy predictor to test against datasets for University of Alabama student enrollment and Lahi crop production [48]. Li et al. (2010) developed a deterministic vector long-term forecasting (DVL) method which was validated through Monte Carlo simulations conducted on authentic datasets [39]. Chen et al. (2011) concerned two factors to construct high-order fuzzy logical relationship (FLR) groups [11]. Cheng et al. (2013) integrated high-order data into the adaptive network-based fuzzy inference system (ANFIS) using the ordered weighted averaging (OWA) operator to predict stock prices in Taiwan [16].

With the soft computation approaches, Lee et al. (2007) implemented the fuzzy logical relationships for temperature prediction by applying the genetic algorithm (GA) technique [37]. Afterward, Yu and Huarng (2010) claimed that artificial neural network technique can be used to solve the time series problems and also integrate with fuzzy sets [61]. Cheng et al. (2010) revised the fuzzy time series model by integrating rough sets and GA to forecast Taiwanese stock prices [14]. Lately, Chang et al. (2011) proposed a modified fuzzy time series model with ANFIS architecture to forecast the Taiwanese stock prices [5]. Hsieh et al. (2011) integrated wavelet transformations and recurrent neural networks (RNN) to forecast international stock market movements [23]. They also used the artificial bee colony algorithm to adjust the parameters the RNN weights and biases. Chen and Kao (2013) integrated particle swarm optimization and support vector machine techniques into their fuzzy time series model which was experimentally shown to outperform existing conventional methods for stock market forecasts [10].

2.2. Entropy-based discretization

Inferring classification rules is a major task in the artificial intelligence and data mining domains [22]. In decision tree classification algorithms, the discretization process is an effective approach to convert continuous attributes or values to if-then rule statements. The discretization process searches the suitable cutting points to segregate the raw dataset into a small set of intervals that feature a high level of class coherence. Discretization performs two primary tasks: determining the number of discrete intervals and determining the width (i.e., boundaries) of the intervals according to the range of values found for a continuous attribute. Typically, a discretization process explicitly contains four steps: (1) sorting the continuous values of the attribute to be discretized, (2) estimating an optimal cut-point for splitting the intervals, (3) relying on the above evaluation function to split the intervals of continuous values, and (4) finally stopping at some point when the default criterion is met. This research adopts the entropy-based discretization approach proposed by Shannon (1949) in the context of information theory and information gain [47]. The entropy-based method will calculate the entropy value for each class of candidate partition and select suitable boundaries for discretization. After that, the entropy will adopt a larger interval covering all values of a feature and continue to partition this bigger interval into smaller subintervals until the specific rule for stopping is met [32].

An object set D is constituted of n classes $(d_1, d_2, \ldots, d_n)$, having its probabilities $p_1, p_2, \ldots, p_n$. Therefore, the entropy of D can be illustrated as follows:

$$Entropy(D) = -\sum_{i=1}^{n} p_i \log_2(p_i) \tag{1}$$

Let an attribute X split D into m disjoint sets $D_1, D_2, \ldots, D_m$, thus the entropy $Entropy(X, D)$ of D partitioned by X is given as follows:

$$Entropy(X, D) = -\sum_{i=1}^{m} \frac{|D_i|}{D} Entropy(D_i) \tag{2}$$

where $|D_i|$ is the amount of tuples in D. Thus, the cutting point for attribute X should provide the attribute value that satisfies the minimum requirement for expected information. Finally, the cutting point selection policy is used repeatedly on each partition until stopping criterion is fulfilled.

Entropy-based discretization has been successfully applied in many domains [32,36,64], but has never been used in forecasting with fuzzy time series models. For this reason, we adopt entropy-based discretization to measure whether the linguistic values approach a steady state belonging to the fuzzy set.

3. Research methodology

This research presents a novel fuzzy time series model based on granular computing with entropy-based and binning-based discretization partitioning to forecast stock prices on the TAIEX, DJIA, S&P 500, and IBOVESPA stock indexes. This section adopted the training dataset from the Taiwan Stock Exchange (from January 1 to October 30, 1999) to illustrate the research procedures.

3.1. Step 1: Define the universe of discourse U

The U is denoted as $[D_{min} - D_1, D_{max} + D_2]$, where D_{min} and D_{max} respectively represent the minimum and maximum stock prices in the historical dataset. In addition, D_1 and D_2 represent the appropriate positive real values.

3.2. Step 2: Segregate U into linguistic intervals with binning-based methods

The discourse universe is segregated into seven linguistic values, based on Miller's (1994) [44] suggestion that the suitable number of classes for human short-term memory effect ranges from five to nine. This step includes two sub-steps used to verify our proposed models:

3.2.1. Partition U into linguistic intervals with equal-width (distance) pre-partitioning

Pre-partitioning based on equal-width (distance) separates all linguistic intervals into the same widths. In the first iteration, the average-based width method is used to divide the universe of discourse U into u_1, u_2, u_3, u_4, u_5, u_6, and u_7 intervals, these intervals are all equal length [6,25]. The resulting linguistic values are then defined and shown in Fig. 1.

3.2.2. (b) Partition U into linguistic intervals with equal-depth (frequency) pre-partitioning

Pre-partitioning based on equal-depth (frequency) creates linguistic intervals with similar frequencies (i.e., the number of contiguous data samples in each interval is approximately the same). The first iteration divides the universe of discourse U into u_1, u_2, u_3, u_4, u_5, u_6, and u_7 intervals, these intervals are all equal depth and linguistic values are defined in Fig. 2. These seven equal depth intervals were obtained by equal-width pre-partitioning in Table 1.

3.3. Step 3: Entropy-based discretization partitioning to define linguistic intervals

Observations can be more reliably fuzzified into optimal linguistic values by adjusting the suitable linguistic interval with granular computing. Thus, in Step 2, entropy-based discretization partitioning can be used to establish optimal definitions for the universe of intervals.

Following the pre-partitioning steps, the training dataset is fuzzified into seven interval values. Each datum is then partitioned using Eqs. (1) and (2) to determine the respective entropy values, thus producing the minimum entropy split-point between $observation_i$ and $observation_{i+1}$ where $observation_i$ is the stock index level of day i.

The training dataset is completely fuzzified into its respective interval values, with linguistic intervals defined in Fig. 3(a), and the membership functions shown in Fig. 3(b). We then sort the data from small to large as shown as Fig. 3(c), and apply the entropy-based discretization method to find the appropriate split-point value by partitioning each datum. Thus, we can obtain the 2nd through 6th split-points respectively and the linguistic intervals of the second iteration as shown in Fig. 4(a), with the membership functions as shown in Fig. 4(b). We next proceed to establish fuzzy sets for observations of the second iteration as in step 4, and obtain the linguistic value corresponding to each datum as shown in Fig. 4(c), which shows that some linguistic values have changed. Finally, through repeated iterations, the linguistic values of each datum eventually cease changing.

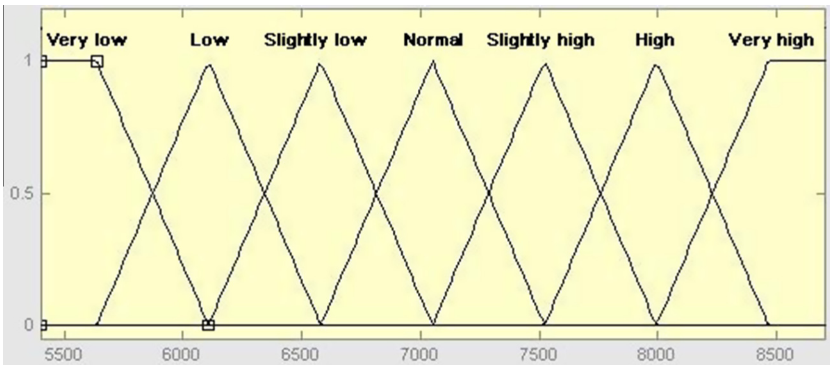

Fig. 1. Membership function of the first iteration (1999) with equal-width pre-partitioning.

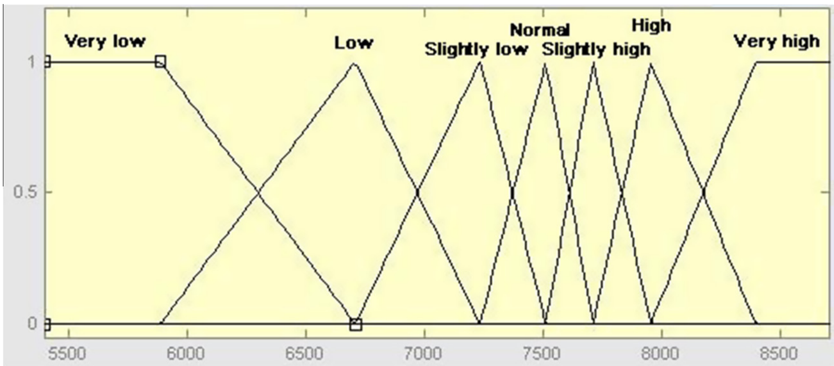

Fig. 2. Membership function of the first iteration (1999) with equal-depth pre-partitioning.

Table 1
Seven intervals of the TAIEX training dataset (1999).

Intervals	1st Iteration	2nd Iteration	3rd Iteration	4th Iteration
u_1	[5400, 5871]	[5400, 5984]	[5400, 5984]	[5400, 5984]
u_2	[5871, 6343]	[5984, 6343]	[5984, 6377]	[5984, 6393]
u_3	[6343, 6814]	[6343, 6823]	[6377, 6823]	[6393, 6881]
u_4	[6814, 7286]	[6823, 7289]	[6823, 7316]	[6881, 7316]
u_5	[7286, 7757]	[7289, 7759]	[7316, 7786]	[7316, 7802]
u_6	[7757, 8229]	[7759, 8265]	[7786, 8265]	[7802, 8265]
u_7	[8229, 8700]	[8265, 8700]	[8265, 8700]	[8265, 8700]

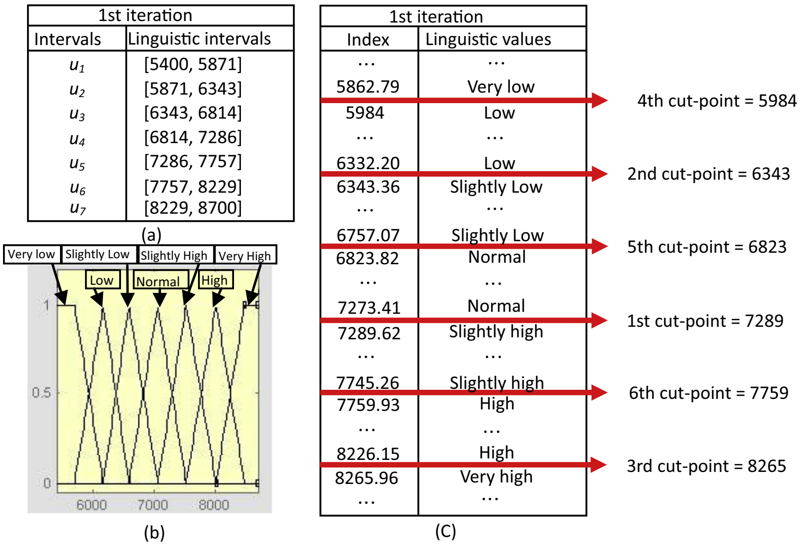

Fig. 3. Entropy-based discretization partitioning of the first iteration (1999).

3.4. Step 4: Define the linguistic values and fuzzify the data

Eq. (3) is used to determine the linguistic values of $A_1, A_2, \ldots, A_n$, where $f_{A(i)}$ denotes the membership function of the fuzzy set $A(i)$, and $f_{A(i)}(u_i)$ denotes the degree of membership of u_i belonging to the fuzzy set $A(i)$, and $f_{A(i)}(u_i) \in [0, 1]$, and $1 \leqslant i \leqslant n$ [6,8].

$$A(i) = f_{A(i)}(u_1)/u_1 + f_{A(i)}(u_2)/u_2 + \cdots + f_{A(i)}(u_n)/u_n \tag{3}$$

Once the optimal linguistic interval settings have been determined, Eq. (3) is used to determine the linguistic values A_1, A_2, …, A_n as follows [6,8,59]:

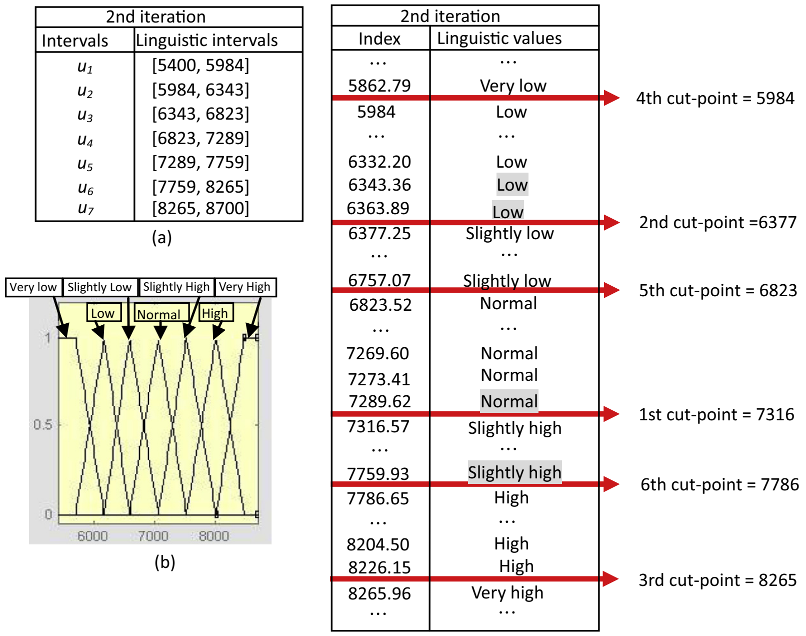

Fig. 4. Entropy-based discretization partitioning of the second iteration (1999).

$$
\begin{aligned}
A_1 &= 1/u_1 + 0.5/u_2 + 0/u_3 + \cdots\cdots + 0/u_{n-2} + 0/u_{n-1} + 0/u_n \\
A_2 &= 0.5/u_1 + 1/u_2 + 0.5/u_3 + \cdots\cdots + 0/u_{n-2} + 0/u_{n-1} + 0/u_n \\
&\cdots \\
&\cdots \\
&\cdots \\
A_{n-1} &= 0/u_1 + 0/u_2 + 0/u_3 + \cdots\cdots + 0.5/u_{n-2} + 1/u_{n-1} + 0.5/u_n \\
A_n &= 0/u_1 + 0/u_2 + 0/u_3 + \cdots\cdots + 0/u_{n-2} + 0.5/u_{n-1} + 1/u_n
\end{aligned}
$$

Table 2 shows seven linguistic values A_1–A_7 obtained by Eq. (3), while Table 2 illustrates several samples of linguistic values for the TAIEX. In addition, Figs. 2 and 3 show the membership functions for the initial iteration of the 1999 training dataset with equal-width pre-partitioning and equal-depth pre-partitioning.

3.5. Step 5: Establish the fuzzy relationships of day i and i + 1

The FLR $A_j \rightarrow A_k$ denotes that "if the stock price of day i is A_j then the day $i+1$ is A_k,". Then A_j is the present stock price while A_k is the stock price of the following day. For example, in the 2nd iteration in Table 3, the FLRs proceed as follows: $A_2 \rightarrow A_2$, $A_2 \rightarrow A_2$, $A_2 \rightarrow A_3$, $A_3 \rightarrow A_3$, $A_3 \rightarrow A_3$, $A_3 \rightarrow A_3$.

3.6. Step 6: Set FLR groups

FLRs with identical LHS values are grouped into ordered FLR groups by grouping their RHSs [59]. A trend-weighted matrix is then generated for all FLRs. For example, as shown in Table 2, the 2nd iteration of FLR groups are $A_2 \rightarrow A_2, A_3$ and $A_2 \rightarrow A_3$, while the 2nd iteration in Table 3 illustrates the trend-weighted matrix for the 1999 training dataset.

Table 2
Linguistic values of the TAIEX training dataset (1999).

Date	Index	1st Iteration	2nd Iteration	3rd Iteration	4th Iteration
1/19	6343.36	Slightly low	Low	Low	Low
1/12	6363.89	Slightly low	Low	Low	Low
1/18	6377.25	Slightly low	Slightly low	Low	Low
3/5	6383.09	Slightly low	Slightly low	Low	Low
3/4	6393.74	Slightly low	Slightly low	Slightly low	Low
3/3	6403.14	Slightly low	Slightly low	Slightly low	Low
1/7	6404.31	Slightly low	Slightly low	Slightly low	Low
1/11	6406.99	Slightly low	Slightly low	Slightly low	Low

Table 3
Trend-weighted matrix for the training dataset (1999).

$P(t-1)$	$P(t)$						
	A_1	A_2	A_3	A_4	A_5	A_6	A_7
A_1	8	1	0	0	0	0	0
A_2	1	31	3	0	0	0	0
A_3	0	2	6	4	0	0	0
A_4	0	0	3	20	5	0	0
A_5	0	0	0	4	63	5	0
A_6	0	0	0	0	4	38	2
A_7	0	0	0	0	0	2	18

3.7. Step 7: Set weighting matrix

The previously-mentioned trend-weighted matrix is used to generate a weighting matrix. The resulting weights are then used to generate a weight matrix $w_i(t)$ and normalized in the follows [59]:

$$w_i(t) = [w'_1, w'_2, \ldots, w'_m] = \left[\frac{w_1}{\sum_{j=1}^{m} w_j}, \frac{w_2}{\sum_{j=1}^{m} w_j}, \ldots, \frac{w_m}{\sum_{j=1}^{m} w_j}\right] \tag{4}$$

where w_i is the weight for fuzzy set A_i, $1 \leqslant i \leqslant k$ and $1 \leqslant j \leqslant m$.

For example, Table 4 shows a weighting matrix produced from Table 3 in which each cell represents the probability of paired FLRs.

3.8. Step 8: Estimating the forecasted values

In step 8, the forecasted values are calculated by both defuzzified matrix and weighted matrix, the equation is defined as follows.

$$F(t) = M_{df}(t-1) \times w_i(t-1) \tag{5}$$

where $M_{df}(t-1)$ denotes the defuzzified matrix. The centroid defuzzification method is then used to derive the weighting matrix $w_i(t-1)$.

3.9. Step 9: Forecasted value modification

To optimize the model's prediction performance, the research applies the adaptive expectation model (AEM) [35] to adjust and increase the forecasting performance. Eq. (6) defines the adaptive expectation model as follows:

$$\text{AEM}(t) = P(t-1) + h * (F(t) - P(t-1)) \tag{6}$$

where AEM(t) denotes the predicted price at time t, while $P(t-1)$ denotes the price at time $t-1$, h is the weight, and $(F(t) - P(t-1))$ denotes the forecasting deviation at time $t-1$. Thus, Step 8 provides the $F(t)$ value, and the forecast is adjusted using the forecasting error $F(t) - P(t-1)$.

4. Experimental results

Large authentic stock price datasets were collected from the TAIEX, DJIA, S&P 500 and IBOVESPA to validate the performance of the proposed models. Section 4.1 below provides a detailed description of the performance assessment criteria and

Table 4
Weighting matrix for the training dataset (1999).

$P(t-1)$	$P(t)$						
	A_1	A_2	A_3	A_4	A_5	A_6	A_7
A_1	0.89	0.11	0	0	0	0	0
A_2	0.02	0.89	0.09	0	0	0	0
A_3	0	0.17	0.5	0.33	0	0	0
A_4	0	0	0.11	0.71	0.18	0	0
A_5	0	0	0	0.06	0.88	0.06	0
A_6	0	0	0	0	0.09	0.86	0.05
A_7	0	0	0	0	0	0.1	0.9

the experimental datasets. Section 4.2 describes three tests and aims to evaluate the proposed models using existing conventional fuzzy time series models, along with the SVM and GARCH models. Finally, Section 4.3 describes the use of paired 2-tailed *t*-tests to contrast with the forecasting results among these models.

4.1. Performance design criteria and experimental period

In this section, the proposed model is evaluated via three experiments. The forecasting comparison and performance evaluation is conducted for a one-step-ahead horizon in all experiments. The first experiment seeks to investigate the proposed model with six fuzzy time series models [5,7,10,16,23,59] and also with support vector regression with three different kernel functions (Polynomial [54], RBF [54], and Puk [53]). The dataset comprises daily stock index of TAIEX, which covered the period from 1997 to 2003. In addition, the TAIEX dataset is divided year by year. The training period is collected from January to October while the testing period is collected from November to December.

Then, the second experiment seeks to validate the accuracy in another financial time series dataset which composed of daily stock index of DJIA from 1997 to 2003. The DJIA is one of the most popular stock indexes based on the market capitalizations of 30 significant stocks traded on the NYSE (New York Stock Exchange) or NASDAQ (National Association of Securities Dealers Automated Quotations). The DJIA dataset is also divided year by year. In addition, the setup method of training period and testing period is the same as first experiment. The comparison models include three fuzzy time series models [8,11,23] and support vector regression with three different kernel functions (Polynomial [54], RBF [54], and Puk [53]).

Finally, the third experiment compares prediction accuracy between the proposed model and modern economic models such as GARCH [3], GJR-GARCH [21], and Fuzzy GJR-GARCH [31] using two datasets composed of daily stock index of S&P 500 and IBOVESPA from 2000 to 2011. The S&P 500 index is based on the market capitalizations of 500 large companies listed on the NYSE or NASDAQ. The IBOVESPA Index is an accumulation index of about 50 largest Brazilian stocks in Brazil. These two datasets are the most popular benchmark for North American economy and South American economy. For these two datasets, the training dataset consists of data from January 2000 through December 2005, and the testing dataset is from January 2006 through September 2011.

The performance evaluation is assessed by comparing the measuring forecast error criteria via the mean squared error (MSE), root of the mean squared error (RMSE), mean absolute error (MAE), mean percentage error (MPE) and Morgan–Granger–Newbold (MGN). These five indicators are always used to evaluate forecasting performance, are widely used in time series models comparisons [3,5,7,8,10,11,16,21,23,31,53,54,59], and are defined by Eqs. (7)–(11):

$$\mathrm{MSE} = \frac{\sum_{i=1}^{n}(forecast(i) - actual(i))^2}{n} \tag{7}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(forecast(i) - actual(i))^2}{n}} \tag{8}$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n}|(forecast(i) - actual(i))|}{n} \tag{9}$$

$$\mathrm{MPE} = \frac{\sum_{i=1}^{n}|(forecast(i) - actual(i))|/actual(i)}{n} \tag{10}$$

$$\mathrm{MGN} = \frac{\hat{\rho}_{xz}}{\left(\frac{1-\hat{\rho}_{xz}^2}{n-1}\right)^{\frac{1}{2}}} \tag{11}$$

where *actual*(i) denotes the actual value at time i, *forecast*(i) denotes the predictive value at time i, n denotes the number of testing period, $\hat{\rho}_{xz}$ denotes the estimated correlation coefficient between $x = \varepsilon 1 + \varepsilon 2$ and $z = \varepsilon 1 - \varepsilon 2$, and $\varepsilon 1$ and $\varepsilon 2$ are the adjusted residuals of the two models.

Table 5
Performance comparisons for the TAIEX with RMSE value (1997–2003).

Models	1997	1998	1999	2000	2001	2002	2003	Average	SD
Chen (1996) [7]	154	134	120	176	148	101	74	129.6	34.4
Yu (2005) [59]	165	164	145	191	167	75	66	139.0	48.7
Chang et al. (2011) [5]	133	117	100	173	119	61	53	108.0	41.5
Hsieh et al. (2011) [23]	126	83	86	135	93	62	59	92.0	29.2
Cheng et al. (2013) [16]	133	115	103	130	120	66	55	103	30.9
Chen and Kao (2013) [10]	125	104	87	125	114	76	54	97	26.7
SVR-Polynomial [54]	233	125	130	190	266	79	109	161.7	69.1
SVR-RBF [54]	301	171	231	208	746	205	200	294.5	203.1
SVR-Puk [53]	282	586	546	1912	337	463	360	640.8	571.4
Proposed model (Width)	115	107	92	115	86	58	41	87.7	28.7
Proposed model (Depth)	115	97	92	118	89	60	41	87.4	28.0

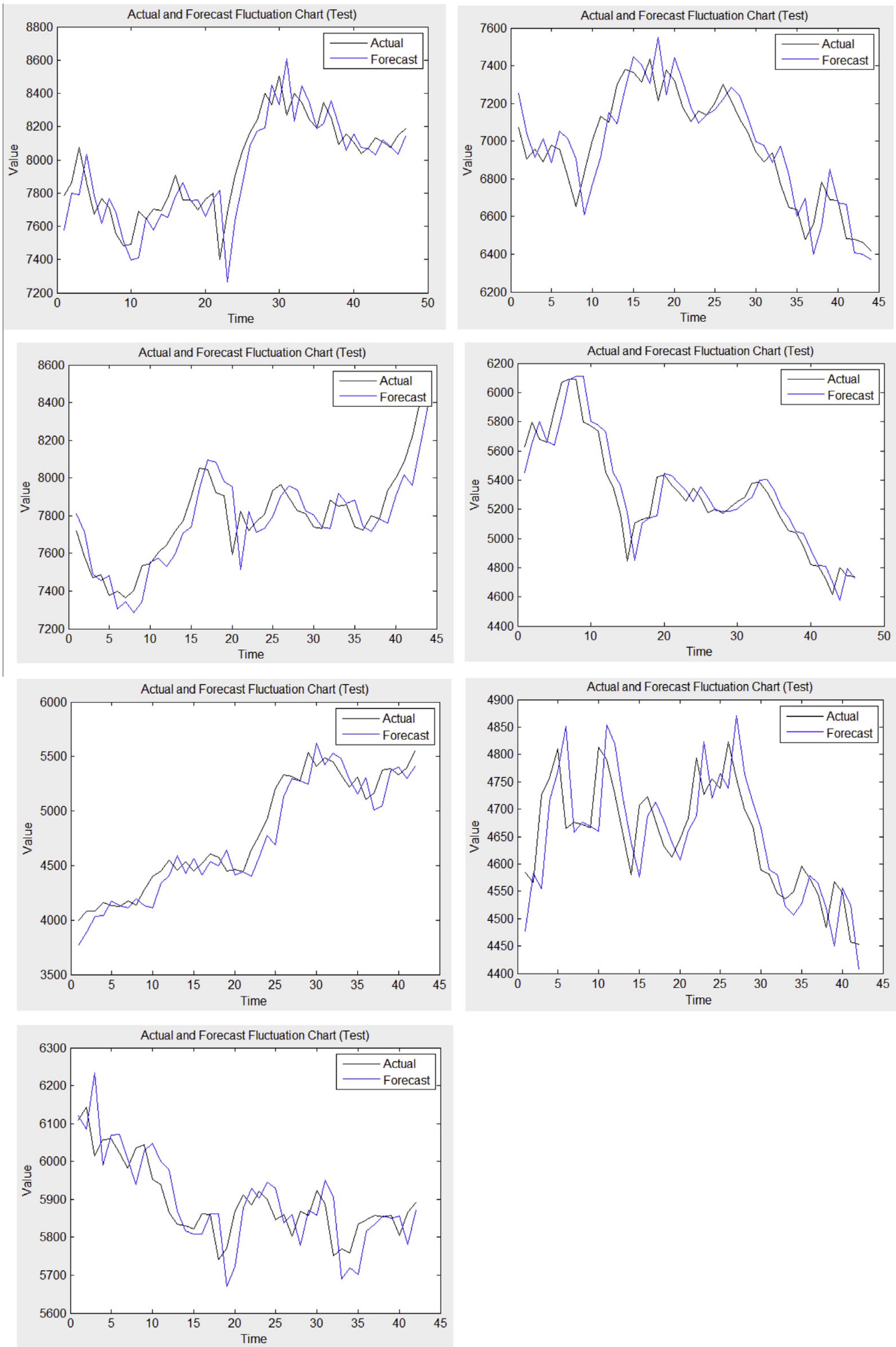

Fig. 5. The stock price fluctuation for TAIEX testing dataset with equal-depth pre-partitioning (1997–2003).

As previously mentioned, the proposed models first decide the numeric range for the universe of discourse. Then, the dataset for each year is calculated, the integer which is nearest to the minimal value will be set as the starting value, and the integer which is nearest to the maximal value will be set as the end value. From the nine testing periods, we generated forecasting values with seven linguistic values $A_1, A_2, \ldots, A_7$, for both equal-width pre-partitioning and equal-depth

Table 6
Performance comparisons for the DJIA with RMSE value (1997–2003).

Models	1997	1998	1999	2000	2001	2002	2003	Average	SD
Chen and Chen (2011) [8]	139	124	115	127	121	74	66	109.4	27.9
Chen et al. (2011) [11]	51	139	113	131	113	65	52	94.8	37.7
Hsieh et al. (2011) [23]	92	90	94	137	104	132	89	105.4	20.5
SVR-Polynomial [54]	346	212	120	178	131	91	139	173.8	85.5
SVR-RBF [54]	503	426	197	315	136	129	94	257.1	160
SVR-Puk [53]	768	477	578	586	1168	679	544	685.7	232.5
Proposed model (Width)	75	81	71	114	75	98	55	81.2	19.2
Proposed model (Depth)	70	81	74	108	87	96	48	80.5	19.3

pre-partitioning in the initial iteration from Step 2. Entropy-based discretization partitioning was used to evaluate the optimal linguistic interval. The resulting observations correspond to fuzzy sets $A_1, A_2, \ldots, A_7$, and cease changing from Steps 3 and 4. However, the forecast results could be calculated by multiplying the corresponding weighting matrix and the deffuzified matrix from Step 5 to Step 8, and by modifying the adapted forecasting stock price by the adaptive expectation model defined in Eq. (6).

4.2. Empirical analysis

The first experiment compared the proposed models with other hybrid fuzzy time series models developed by Chen (1996) [7], Yu (2005) [59], Chang et al. (2011) [5], Hsieh et al. (2011) [23], Cheng et al. (2013) [16], and Chen and Kao (2013) [10], and the RMSE results as shown in Table 5. In addition, this research also compared them to the support vector regression (SVR) models including the Polynomial kernel function [54], RBF kernel function [54], and Puk kernel function [53]. The experimental dataset consisted of daily data from TAIEX. In nearly all seven testing periods, the proposed models produce the smallest RMSE values for both equal-width and equal-depth pre-partitioning, with the exceptions of 1998 and 1999. With an equal-depth pre-partitioning of 87.4, the proposed model produces the smallest average RMSE value, while that using equal-width pre-partitioning produced an average RMSE value of 87.7. These results indicate that applying the entropy-based discretization partitioning approach to the equal-depth model effectively reduces forecasting errors. The proposed model using equal-depth pre-partitioning has the smallest indicator of standard deviation (SD) consistency at 28.0, meaning it provides the most consistent performance, thus the experimental results of the proposed models is obviously more consistent than provided by the hybrid fuzzy time series models of Chen (1996) [7], Yu (2005) [59], Chang et al. (2011) [5], Hsieh et al. (2011) [23], Cheng et al. (2013) [16], and Chen and Kao (2013) [10]. The proposed models also outperform the SVR models including Polynomial [54], RBF [54], and Puk [53]. In Fig. 5, the black line expresses the actual TAIEX stock price, and the blue line expresses the forecasting TAIEX stock price of entropy-based approach with equal-depth pre-partitioning. The *Y*-axis represents the TAIEX stock price and the *X*-axis represents the number of trading days of testing period. During these seven periods, it is found the proposed approach has smoothly curve and close in the actual TAIEX stock price both in uptrend and downtrend of stock price changes. Therefore, the proposed models outperformed all of the other models in terms of average RMSE value consistency and the results were listed in Table 5.

The second experiment used a dataset comprising daily data from DJIA. This experiment compared the proposed models to both the hybrid models developed by Chen and Chen (2011) [8], Chen et al. (2011) [11], and Hsieh et al. (2011) [23], and three SVR models with Polynomial [54], RBF [54], and Puk [53] kernel functions. As shown in Table 6, for the seven testing periods, the proposed models (both equal-width and equal-depth pre-partitioning) exhibit the lowest RMSE values. The proposed model adopting equal-depth pre-partitioning had the highest average RMSE value of 80.5, while the proposed model adopting equal-width pre-partitioning had the smallest SD indicator at 19.2, followed by the equal-depth approach. In tests based on the DJIA dataset, the proposed models clearly outperformed the models developed by Chen and Chen (2011) [8], Chen et al. (2011) [11], Hsieh et al. (2011) [23], SVR models with Polynomial [54], RBF [54], and Puk [53] kernel functions in terms of prediction accuracy. In Fig. 6, the black line expresses the actual DJIA stock price, and the blue line expresses the forecasting DJIA stock price of entropy-based approach with equal-depth pre-partitioning. The Y-axis represents the DJIA stock price and the X-axis represents the number of trading days of testing period. The experimental results illustrated that the proposed approach aligned closely with the actual DJIA stock prices. Meanwhile, the models with RMSE value and SD also outperformed the other models, as shown in Table 6.

The third experiment used a dataset consisting of daily closing prices of the S&P 500 and IBOVESPA from 2000 to 2011. The parameter estimates for the training dataset associated with the Fuzzy GJR-GARCH, GARCH(1,1), and GJR-GARCH(1,1) models were referenced from Maciel's study (2012) [42]. Table 7 shows the performance of the evaluated models to predict volatility on the S&P 500 and IBOVESPA stock indexes. The proposed models outperformed the other models [3,21,31] in terms of estimating forecast errors. The proposed models and the Fuzzy GJR-GARCH model demonstrated similar results, but the proposed models had slightly fewer errors than the Fuzzy GJR-GARCH model for the IBOVESPA index.

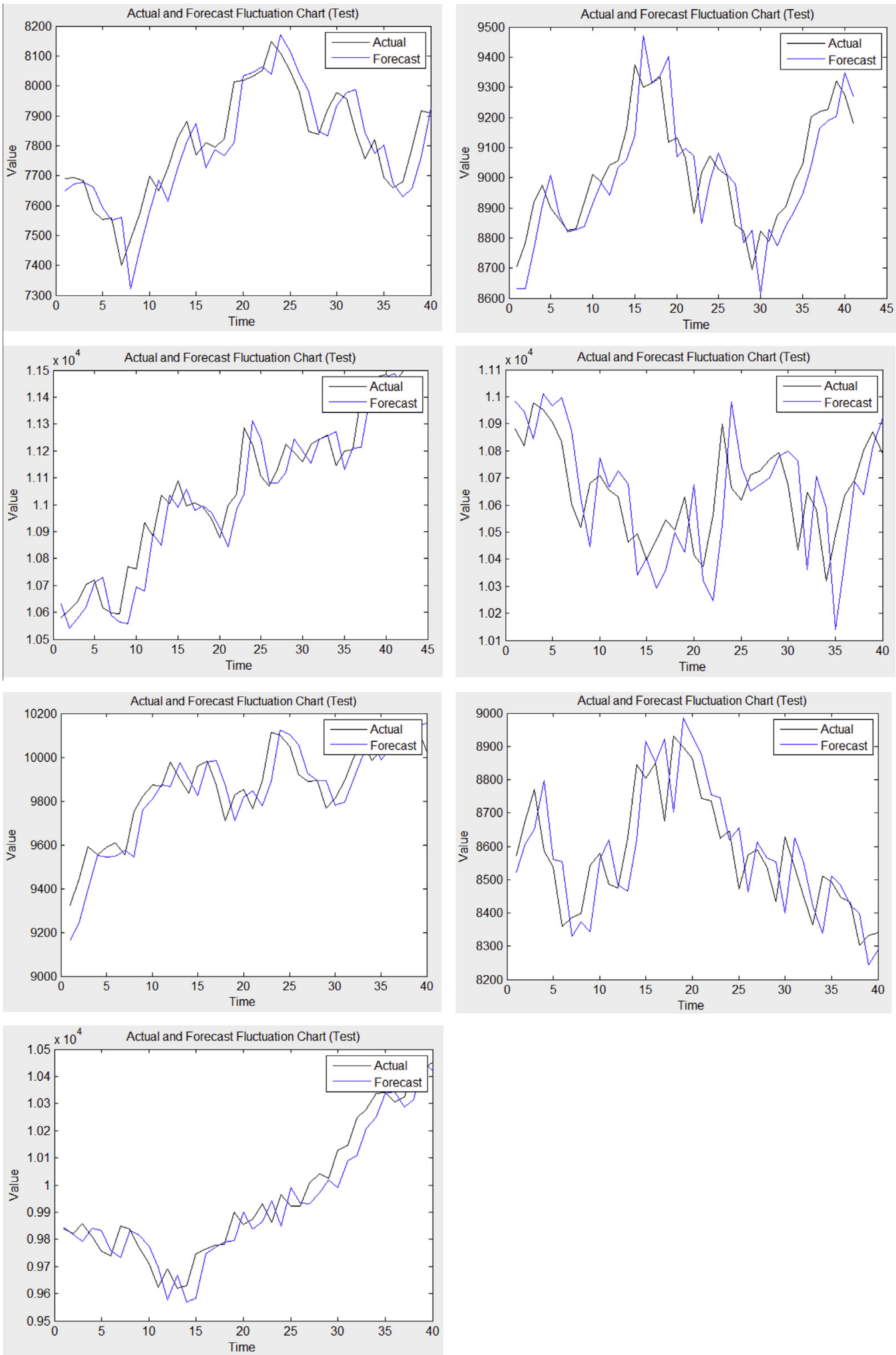

Fig. 6. The stock price fluctuation for DJIA testing dataset with equal-depth pre-partitioning (1997–2003).

Table 7
Performance comparisons for one-step ahead (2000–2011).

	RMSE	MSE	MAE	MPE
Models for S&P 500				
Fuzzy GJR-GARCH [31]	0.4453	0.1983	0.3652	0.3795
GARCH [3]	0.7552	0.5704	0.7076	0.8366
GJR-GARCH [21]	0.7641	0.5839	0.7298	0.8420
Proposed model (Width)	0.3377	0.1140	0.2197	0.2614
Proposed model (Depth)	0.3070	0.0942	0.1676	0.1986
Models for IBOVESPA				
Fuzzy GJR-GARCH [31]	0.7809	0.6099	0.7912	0.2852
GARCH [3]	1.2036	1.4487	1.2403	0.6723
GJR-GARCH [21]	1.1928	1.4230	1.1955	0.6511
Proposed model (Width)	0.5833	0.3402	0.6082	0.1591
Proposed model (Depth)	0.5821	0.3388	0.6058	0.1583

Table 8
Statistical tests for TAIEX and DJIA datasets.

Compared Models	Paired 2-tailed *t*-tests	
	Proposed model (Width)	Proposed model (Depth)
TAIEX dataset	*p*-Value/diff.	*p*-Value/diff.
Chen (1996) [7]	0.029**/−41.90	0.027**/ -42.20
Yu (2005) [59]	0.034**/−51.30	0.032**/ -51.60
Chang et al. (2011) [5]	0.309/−20.30	0.299/−20.60
Hsieh et al. (2011) [23]	0.787/−4.30	0.770/−4.60
Cheng et al. (2013) [16]	0.352/−15.30	0.339/−15.60
Chen and Kao (2013) [10]	0.507/−9.30	0.490/−9.60
SVR-Polynomial [54]	0.023**/−74.00	0.022**/−74.30
SVR-RBF [54]	0.021**/−206.80	0.020**/207.10
SVR-Puk [53]	0.025**/−553.10	0.025**/−553.40
DJIA dataset	*p*-value/diff.	*p*-value/diff.
Hsieh et al. (2011) [23]	0.042**/−24.20	0.038**/−24.90
Chen and Chen (2011) [8]	0.049**/−28.20	0.045**/−28.90
Chen et al. (2011) [11]	0.414/−13.60	0.391/−14.30
SVR-Polynomial [54]	0.016**/−92.60	0.016**/−93.30
SVR-RBF [54]	0.014**/−175.90	0.013**/−176.60
SVR-Puk [53]	0.000***/−604.50	0.000***/−605.20

[a]Diff. denotes the difference between the proposed models and compared models in average.
[b]Confidence level $\alpha = 0.05$.
** Denotes the significance at 5%.
*** Denotes the significance at 1%.

4.3. Statistical analysis

In the first experiment, statistical tests were used to examine the prediction performance between the proposed models with those put forward by Chen (1996) [7], Yu (2005) [59], Chang et al. (2011) [5], Hsieh et al. (2011) [23], Cheng et al. (2013) [16], and Chen and Kao (2013) [10], as well as the SVR-Polynomial [54], SVR-RBF [54], and SVR-Puk [53] models for the TAIEX validation sets from 1997 to 2003. Table 8 shows the results of the paired two-tailed *t*-tests which indicate that the proposed models using equal-width and equal-depth pre-partitioning outperformed five of the seven models at an almost 5% statistical significance level, with the exceptions of Chang et al. (2011) [5], Hsieh et al. (2011) [23]. The paired two-tailed *t*-test shows that the predictions of the proposed models and those of hang et al. (2011) [5] and Hsieh et al. (2011) [23] are equally accurate. However, *t*-test results are easily influenced by sample size, so we then calculated the difference between the proposed models and compared models on average with *Diff* values in Table 8. The proposed models were found to have smaller average RMSE values and higher *Diff* values than either of these two models. Statistical analysis found the proposed models provide better forecasting performance than the other hybrid time series models.

In the second experiment, statistical tests were used to investigate the prediction accuracy of the proposed models to that the models developed by Chen and Chen (2011) [8], Chen et al. (2011) [11], and Hsieh et al. (2011) [23], as well as the SVR-Polynomial [54], SVR-RBF [54], and SVR-Puk [53] models for the DJIA validation sets from 1997 to 2003. Table 8 presents the results of the paired two-tailed *t*-tests. The paired *t*-test results were below 5%, indicating that the proposed models outperform nearly all other models in terms of prediction accuracy. However, the paired two-tailed *t*-test suggests that the predictions of the proposed models and those of Chen et al. (2011) [11] are equally accurate. In fact, the proposed models

Table 9
MGN Statistical test for S&P 500 and IBOVESPA datasets.

Compared models	Morgan–Granger–Newbold statistical test	
	lProposed model (Width)	lProposed model (Depth)
S&P 500 Dataset	*MGN* value/*p*-value	*MGN* value/*p*-value
Fuzzy GJR-GARCH [31]	8.2497/0.0023***	7.6646/0.0052***
GARCH [3]	7.4701/0.0001***	8.0263/0.0011***
GJR-GARCH [21]	10.0673/0.0000***	11.0793/0.0000***
IBOVESPA Dataset	*MGN* value/*p*-Value	*MGN* value/*p*-Value
Fuzzy GJR-GARCH [31]	6.1756/0.0071***	6.1784/0.0070***
GARCH [3]	1.3825/0.0123**	1.3828/0.0122**
GJR-GARCH [21]	1.3033/0.0135**	1.3036/0.0135**

[a]Confidence level $\alpha = 0.05$.
** Denotes the significance at 5%.
*** Denotes the significance at 1%.

still have a smaller average RMSE value and higher *Diff* values than those of Chen et al.'s (2011) [11] model. This demonstrates that the scalability and robust of the proposed models are sufficient to deal with financial time series datasets and provide reliable predictions.

Table 9 shows the MGN statistical test results for the third experiment. In Table 9, the proposed models outperform the other models with significant *p*-values. Therefore, the MGN statistical tests support that the proposed models provide superior performance for forecasting S&P 500 and IBOVESPA datasets than Fuzzy GJR-GARCH model [31], traditional GARCH model [3], and GJR-GARCH model [21]. The combined results of these three experiments indicate that the proposed models are highly capable of forecasting real stock market pricing.

5. Conclusions

In this research, the novel hybrid fuzzy time series model has been proposed based on granular computing approach to predict stock prices. The proposed model has significant performance than existing fuzzy time series models, hybrid intelligent forecasting models and traditional economic forecasting models for all testing datasets (TAIEX, DJIA, S&P 500, and IBOVESPA. Our findings investigate several scientific contributions as follows.

First, this research proposes the fuzzy time series model based on the granular computing approach. It adopts the entropy-based discretization method to regulate the interval lengths during the iteration process. Consequently, it can impartially decide the reasonable intervals. In addition, the proposed model also uses the binning-based partition to determine reasonable interval lengths by partitioning the universe of discourse and related linguistic values of each datum to change through repeated iterations.

Second, the experimental results illustrate the entropy-based discretization partitioning can be an appropriate method than existing fuzzy time series models, and support the use of vector regression for developing a model for forecasting stock prices. As shown in Tables 5 and 6, the proposed model has lower average RMSEs than the current methods [5,7,8,10,11,16,23,53,54,59]. This conclusion also supports the findings of Lu et al. (2009) [40] and Lu and Wu (2011) [41] in that the SVR models did not obtain a higher accuracy rate than intelligent computation methods and can use only the hybrid approach to enhance prediction performance.

Finally, the empirical results also show that stock price forecasting models developed using entropy-based discretization partitioning can obtain lower RMSEs than conventional GARCH, GJR-GARCH, and Fuzzy GARCH models. As shown in Table 7, the proposed method provides improved prediction results than the models proposed in [3,21,31] for the S&P 500 and IBOVESPA indexes. Although the Fuzzy GARCH model does not outperform the proposed time series model, it is better able handle more input variables in time series datasets [29,30]. In this research, the only input variable was the daily closing price and no other technical analysis indicators (e.g., trading volume and financial reports) were included. Therefore, we recommend further investigation into the differences between these two models.

This research used datasets from the TAIEX, DJIA, S&P 500, and IBOVESPA, and further validation is needed using data from other real world stock markets. Second, this research discusses the first-order fuzzy time series models, and future research could attempt to create higher-order fuzzy time series models to subjectively account for the multi-factor causality inherent in stock indices. The proposed approach could also be extended by adopting other important factors as the variables, including trading volume, financial reports, and other technical analysis indicators.

Acknowledgments

The authors thank the support of National Scientific Council (NSC) of the Republic of China (ROC) to this work under Grant No. NSC-101-2410-H-025-004-MY2 and MOST-103-2410-H-025-022-MY2. The authors also gratefully acknowledge the Editor, Associate Editor, and anonymous reviewers for their valuable comments and constructive suggestions.

Appendix A. Definitions of abbreviations

Adaptive network-based fuzzy inference system, ANFIS
Autoregressive conditional heteroscedastic, ARCH
Autoregressive moving average, ARMA
Dow-Jones Industrial Average, DJIA
Fuzzy Logical Relationship, FLR
Generalized ARCH, GARCH
Glosten–Jagannathan–Runkle GARCH, GJR-GARCH
Left-hand side, LHS
Mean absolute error, MAE
Mean percentage error, MPE
Mean squared error, MSE
Morgan–Granger–Newbold, MGN
Ordered weighted averaging, OWA
Right-hand side, RHS
Root of the mean squared error, RMSE
Recurrent neural networks, RNN
Support vector regression, SVR
Taiwan Stock Exchange Capitalization Weighted Stock Index, TAIEX

References

[1] E.I. Altman, Financial ratios discriminant analysis and the prediction of corporate bankruptcy, J. Finance 23 (1968) 589–609.
[2] W.H. Beaver, Financial ratios as predictors of failure, J. Account. Res. 4 (1966) 71–111.
[3] T. Bollerslev, Generalized autoregressive conditional heteroscedasticity, J. Econ. 31 (1986) 307–327.
[4] G. Box, G. Jenkins, Time Series Analysis: Forecasting and Control, Holden-Day, San Francisco, 1976.
[5] J.R. Chang, L.Y. Wei, C.H. Cheng, A hybrid ANFIS model based on AR and volatility for TAIEX Forecasting, Appl. Soft Comput. 11 (2011) 1388–1395.
[6] M.Y. Chen, B.T. Chen, Online fuzzy time series analysis based on entropy discretization and a fast Fourier transform, Appl. Soft Comput. 14 (B) (2014) 156–166.
[7] S.M. Chen, Forecasting enrollments based on fuzzy time series, Fuzzy Sets Syst. 81 (1996) 311–319.
[8] S.M. Chen, C.D. Chen, TAIEX forecasting based on fuzzy time series and fuzzy variation groups, IEEE Trans. Fuzzy Syst. 19 (1) (2011) 1–12.
[9] S.M. Chen, N.Y. Chung, Forecasting enrollments of students by using fuzzy time series and genetic algorithms, Inform. Manage. Sci. 17 (3) (2006) 1–17.
[10] S.M. Chen, P.Y. Kao, TAIEX forecasting based on fuzzy time series, particle swarm optimization techniques and support vector machines, Inform. Sci. 247 (2013) 62–71.
[11] S.M. Chen, G.M.T. Manalu, S.C. Shih, T.W. Sheu, H.C. Liu, A new method for fuzzy forecasting based on two-factors high-order fuzzy-trend logical relationship groups and particle swarm optimization techniques, in: Proceedings of 2011 IEEE International Conference on Systems, Man, and Cybernetics, Anchorage, Alaska, 2011, pp. 2301–2306.
[12] C.H. Cheng, T.L. Chen, C.H. Chiang, Trend-weighted fuzzy time-series model for TAIEX forecasting, Lect. Notes Comput. Sci. 4234 (2006) 469–477.
[13] C.H. Cheng, T.L. Chen, H.J. Teoh, C.H. Chiang, Fuzzy time-series based on adaptive expectation model for TAIEX forecasting, Expert Syst. Appl. 34 (2008) 1126–1132.
[14] C.H. Cheng, T.L. Chen, L.Y. Wei, A hybrid model based in rough sets theory and genetic algorithms for stock price forecasting, Inform. Sci. 180 (2010) 1610–1629.
[15] C.H. Cheng, J.W. Wang, C.H. Li, Forecasting the number of outpatient visits using a new fuzzy time series based on weighted-transitional matrix, Expert Syst. Appl. 34 (2008) 2568–2575.
[16] C.H. Cheng, L.Y. Wei, J.W. Liu, T.L. Chen, OWA-based ANFIS model for TAIEX forecasting, Econ. Model. 30 (2013) 442–448.
[17] R. Dash, R.L. Paramguru, R. Dash, Comparative analysis of supervised and unsupervised discretization techniques, Int. J. Adv. Sci. Technol. 2 (3) (2011) 29–37.
[18] J. Dougherty, R. Kohavi, M. Sahami, Supervised and unsupervised discretization of continuous features, in: A. Prieditis, S.J. Russell (Eds.), Morgan Kaufmann, 1995, pp. 194–202.
[19] R.F. Engle, Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, Econometrica 50 (1982) 987–1007.
[20] F. Gaxiola, P. Melin, F. Valdez, O. Castillo, Interval type-2 fuzzy weight adjustment for backpropagation neural networks with application in time series prediction, Inform. Sci. 260 (2014) 1–14.
[21] L.R. Glosten, R. Jagannathan, D.E. Runkle, On the relation between the expected value and the volatility of the nominal excess return on stocks, J. Finance 48 (1993) 1779–1801.
[22] J. Han, K. Micheline, Data Mining Concepts and Techniques, Morgan Kauffman, 2001.
[23] T.J. Hsieh, H.F. Hsiao, W.C. Yeh, Forecasting stock markets using wavelet transforms and recurrent neural networks: an integrated system based on artificial bee colony algorithm, Appl. Soft Comput. 11 (2011) 2510–2525.
[24] K.H. Huarng, T.H.K. Yu, Ratio-based lengths of intervals to improve fuzzy time series forecasting, IEEE Trans. Syst., Man, Cybernet. – Part B: Cybernet. 36 (2006) 328–340.
[25] K.H. Huarng, Effective lengths of intervals to improve forecasting in fuzzy time series, Fuzzy Sets Syst. 123 (3) (2001) 387–394.
[26] K.H. Huarng, T.H.K. Yu, A type 2 fuzzy time series model for stock index forecasting, Physica A 353 (2005) 445–462.
[27] K.H. Huarng, T.H.K. Yu, The application of neural networks to forecast fuzzy time series, Physica A 363 (2006) 481–491.
[28] K.H. Huarng, T.H.K. Yu, Y.W. Hsu, A multivariate heuristic model for fuzzy time-series forecasting, IEEE Trans. Syst., Man, Cybernet., Part B: Cybernet. 37 (4) (2007) 836–846.
[29] J.C. Hung, Adaptive Fuzzy-GARCH model applied to forecasting the volatility of stock markets using particle swarm optimization, Inform. Sci. 181 (2011) 4673–4683.
[30] J.C. Hung, A fuzzy asymmetric GARCH model applied to stock markets, Inform. Sci. 179 (22) (2009) 3930–3943.

[31] J.C. Hung, Applying a combined fuzzy systems and GARCH model to adaptively forecast stock market volatility, Appl. Soft Comput. 11 (5) (2011) 3938–3945.
[32] D. Janssens, T. Brijs, K. Vanhoof, G. Wets, Evaluating the performance of cost-based discretization versus entropy- and error-based discretization, Comput. Oper. Res. 33 (2006) 3107–3123.
[33] T.A. Jilani, S.M.A. Burney, M-factor high order fuzzy time series forecasting for road accident data: analysis and design of intelligent systems using soft computing techniques, Adv. Soft Comput. 41 (2007) 246–254.
[34] S.M. Kendall, K. Ord, Time Series, third ed., Oxford University Press, New York, 1990.
[35] J. Kmenta, Elements of Econometrics, MacMillan, 1986.
[36] D.D. Le, S. Satoh, Ent-boost: boosting using entropy measures for robust object detection, Pattern Recogn. Lett. 28 (2007) 1083–1090.
[37] L.W. Lee, L.H. Wang, S.M. Chen, Temperature prediction and TAIEX forecasting based on fuzzy logical relationships and genetic algorithms, Expert Syst. Appl. 33 (2007) 539–550.
[38] S.T. Li, Y.C. Cheng, Deterministic fuzzy time series model for forecasting enrollments, Comput. Math. Appl. 53 (2007) 1904–1920.
[39] S.T. Li, S.C. Kuo, Y.C. Cheng, C.C. Chen, Deterministic vector long-term forecasting for fuzzy time series, Fuzzy Sets Syst. 161 (2010) 1852–1870.
[40] C.J. Lu, T.S. Lee, C.C. Chiu, Financial time series forecasting using independent component analysis and support vector regression, Decis. Support Syst. 47 (2009) 115–125.
[41] C.J. Lu, J.Y. Wu, An efficient CMAC neural network for stock index forecasting, Expert Syst. Appl. 38 (2011) 15194–15201.
[42] L. Maciel, A hybrid fuzzy GJR-GARCH modeling approach for stock market volatility forecasting, Braz. Rev. Finance 10 (3) (2012) 337–367.
[43] A. Marszałek, T. Burczyński, Modeling and forecasting financial time series with ordered fuzzy candlesticks, Infom. Sci. 273 (2014) 144–155.
[44] G.A. Miller, The magical number seven, plus or minus two: some limits on our capacity of processing information, Psychol. Rev. 101 (1994) 343–352.
[45] C.M. Own, P.T. Yu, Forecasting fuzzy time series on a heuristic high-order model, Cybernet. Syst.: An Int. J. 36 (2005) 705–717.
[46] W. Pedrycz, Granular Computing: Analysis and Design of Intelligent Systems, CRC Press/Francis Taylor, Boca Raton, 2013.
[47] C.E. Shannon, W. Weaver, The Mathematical Theory of Information, University of Illinois Press, Urbana, IL, 1949.
[48] S.R. Singh, A computational method of forecasting based on high-order fuzzy time series, Expert Syst. Appl. 36 (2009) 10551–10559.
[49] Q. Song, B.S. Chissom, Forecasting enrollments with fuzzy time series – Part 1, Fuzzy Sets Syst. 54 (1993) 1–9.
[50] Q. Song, B.S. Chissom, Forecasting enrollments with fuzzy time series – Part 2, Fuzzy Sets Syst. 62 (1994) 1–8.
[51] Q. Song, B.S. Chissom, Fuzzy time series and its models, Fuzzy Sets Syst. 54 (1993) 269–277.
[52] J. Sullivan, W.H. Woodall, A comparison of fuzzy forecasting and Markov modeling, Fuzzy Sets Syst. 64 (1994) 279–293.
[53] B. Üstün, W.J. Melssen, LMC Buydens, Facilitating the application of support vector regression by using a universal Pearson VII function based kernel, Chemometr. Intell. Lab. Syst. 81 (1) (2006) 29–40.
[54] V.N. Vapnik, The Nature of Statistical Learning Theory, Springer-Verlag, New York, 1995.
[55] C.H. Wang, L.C. Hsu, Constructing and applying an improved fuzzy time series model: taking the tourism industry for example, Expert Syst. Appl. 34 (4) (2008) 2732–2738.
[56] N.Y. Wang, S.M. Chen, Temperature prediction and TAIFEX forecasting based on automatic clustering techniques and two-factors high-order fuzzy time series, Expert Syst. Appl. 36 (2:1) (2009) 2143–2154.
[57] R.C. West, A factor-analytic approach to bank condition, J. Bank. Finance 9 (1985) 253–266.
[58] Y. Yao, Granular computing for data mining, in: Proceedings of SPIE Conference on Data Mining, Intrusion Detection, Information Assurance, and Data Networks Security, Paper ID: 624105, 2006.
[59] T.H.K. Yu, Weighted fuzzy time series models for TAIEX forecasting, Physica A 349 (2005) 609–624.
[60] T.H.K. Yu, K.H. Huarng, A bivariate fuzzy time series model to forecast the TAIEX, Expert Syst. Appl. 34 (2008) 2945–2952.
[61] T.H.K. Yu, K.H. Huarng, A neural network-based fuzzy time series model to improve forecasting, Expert Syst. Appl. 37 (2010) 3366–3372.
[62] L.A. Zadeh, The concept of a linguistic variable and its application to approximation reasoning – Part I, Inform. Sci. 8 (1975) 199–249.
[63] L.A. Zadeh, Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets Syst. 19 (1997) 111–127.
[64] M. Zarinbal, M.H. Fazel Zarandi, I.B. Turksen, Relative entropy fuzzy c-means clustering, Inform. Sci. 260 (2014) 74–97.
[65] J. Zhao, K. Liu, W. Wang, Y. Liu, Adaptive fuzzy clustering based anomaly data detection in energy system of steel industry, Inform. Sci. 259 (2014) 335–345.