Available online at www.sciencedirect.com
ScienceDirect
Physica A 380 (2007) 377–390
PHYSICA A
www.elsevier.com/locate/physa

Fuzzy time-series based on Fibonacci sequence for stock price forecasting

Tai-Liang Chen[a,*], Ching-Hsue Cheng[a], Hia Jong Teoh[a,b]

[a]Department of Information Management, National Yunlin University of Science and Technology, 123, Section 3, University Road, Touliu, Yunlin 640, Taiwan, ROC
[b]Department of Accounting Information, Ling Tung University, 1, Ling Tung Road, Nantun, Taichung 408, Taiwan, ROC

Received 31 October 2006; received in revised form 9 February 2007
Available online 6 March 2007

Abstract

Time-series models have been utilized to make reasonably accurate predictions in the areas of stock price movements, academic enrollments, weather, etc. For promoting the forecasting performance of fuzzy time-series models, this paper proposes a new model, which incorporates the concept of the Fibonacci sequence, the framework of Song and Chissom's model and the weighted method of Yu's model. This paper employs a 5-year period TSMC (Taiwan Semiconductor Manufacturing Company) stock price data and a 13-year period of TAIEX (Taiwan Stock Exchange Capitalization Weighted Stock Index) stock index data as experimental datasets. By comparing our forecasting performances with Chen's (Forecasting enrollments based on fuzzy time-series. Fuzzy Sets Syst. 81 (1996) 311–319), Yu's (Weighted fuzzy time-series models for TAIEX forecasting. Physica A 349 (2004) 609–624) and Huarng's (The application of neural networks to forecast fuzzy time series. Physica A 336 (2006) 481–491) models, we conclude that the proposed model surpasses in accuracy these conventional fuzzy time-series models.

Keywords: Fuzzy time series; Fibonacci sequence; Stock price forecasting; Fuzzy linguistic variable

1. Introduction

Time-series models have utilized the fuzzy theory to solve various domains forecasting problems, such as university enrollment forecasting [1–8], financial forecasting [9–12] and temperature forecasting. Especially in the area of stock price forecasting, fuzzy time-series models are often employed [9–12]. As Dourra and Pepe [13] note, it is common practice to "deploy fuzzy logic engineering tools in the finance arena, specifically in the technical analysis field, since technical analysis theory consists of indicators used by experts to evaluate stock price" [13,14].

In stock technical analysis fields, Elliott [42] proposed the Elliott wave principle, which has been playing an important role in stock analysis for more than six decades [15,16]. The theory is closely related to stock price time-series because it applies the Fibonacci sequence to predict the *timing* of stock price fluctuations [17].

*Corresponding author. Tel.: +886 920975168; fax: +886 4 26656058.
E-mail address: g9320817@yuntech.edu.tw (T.-L. Chen).

0378-4371/$ - see front matter
doi:10.1016/j.physa.2007.02.084

Therefore, we argue that applying the sequence in time-series models is a theoretically sound approach whenever the forecasted targets are financial materials.

Lastly, based on the drawbacks of previous models, we propose a new fuzzy time-series model and recommend three refinements to the forecasting processes. In empirical analysis, we employ two experimental datasets to test the proposed model: (1) a 5-year period of TSMC (2000–2004) stock price data and (2) a 13-year period of TAIEX stock index data. And we use three fuzzy time-series models, Chen's [3], Yu's [10] and Huarng's [24] models, as comparison models to verify our model. The performance comparisons show that our model outperforms these conventional fuzzy time-series models.

The remaining content of this paper is organized as follows: Section 2 introduces the related literature of fuzzy time-series models; Section 3 proposes a new fuzzy time-series model and algorithm; Section 4 evaluates the performance of the proposed model; Section 5 includes findings and discussions; and Section 6 provides a conclusion.

2. Related works

This section reviews fuzzy time-series models, the Fibonacci sequence and the Elliott wave principle.

2.1. Fuzzy time-series models

Fuzzy theory was originally developed to deal with problems involving linguistic terms [18–20]. Time-series models had failed to consider the application of this theory until fuzzy time-series was defined by Song and Chissom [1,2]. They proposed the definitions of fuzzy time-series and methods to model fuzzy relationships among observations. The framework of Song and Chissom's model includes six steps: (1) define and partition the universe of discourse; (2) define fuzzy sets for the observations; (3) fuzzify the observations, (4) establish the fuzzy relationship, R, (5) forecast, and (6) defuzzify the forecasting results [1,2].

In the evolution of fuzzy time-series models, Chen [3] proposed another method [3] to apply simplified arithmetic operations in forecasting algorithms rather than the complicated maximum–minimum composition operations, which are presented in Song and Chissom's models. In subsequent research, Chen proposed several methods, such as high-order fuzzy relationships and genetic algorithms to improve his initial model [4–7].

Additionally, Huarng [21] pointed out that the length of intervals affected the forecasting accuracy in fuzzy time-series and proposed a method with distribution-based length and 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 distribution-based and average-based lengths could improve forecasting accuracy [8,21]. 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 [22], establishing linguistic values and dividing intervals would be a trade-off between human recognition and forecasting accuracy [22].

Besides, Yu [10] proposed a weighted model to tackle two issues, recurrence and weighting, in fuzzy time-series forecasting and concluded that the weighted model outperforms one of the conventional fuzzy time-series models [23]. The researcher argued that recurrent fuzzy relationships should be considered in forecasting, and he recommended that different weights be assigned to various fuzzy relationships. Based on statistic theory, we agree with Yu on the weighted approach.

Among recent research, genetic algorithms and neural networks algorithms are usually applied and efficiently improve the forecasting accuracy of fuzzy time-series models [6,24]. However, we believe that there are other reasonable approaches, such as stock analysis theory or techniques of application in time-series models, whenever the forecasted targets are financial materials [11].

2.2. The Fibonacci sequence and Elliott wave principle

The Fibonacci sequence first appeared as the solution to a problem in the Liber Abaci, a book written in 1202 by Leonardo Fibonacci of Pisa to introduce the Hindu–Arabic numerals used today to a Europe still using cumbersome Roman numerals. The original problem in the Liber Abaci asked how many pairs of rabbits can be generated from a single pair, if each month each mature pair brings forth a new pair, which, from the second month, becomes productive [25,44].

In mathematics, the Fibonacci numbers are the sequence of numbers which is defined by

$$F_n = \left\{ \begin{array}{ll} 0 & \text{if } n = 0 \\ 1 & \text{if } n = 1 \\ F_{n-1} + F_{n-2} & \text{if } n > 1 \end{array} \right\}_{n=1.}^{\infty} \tag{1}$$

The Fibonacci sequence also can be expressed by the general second-order linear recurrence equation (where A and B are constants with arbitrary x_1 and x_2), which is defined in Eq. (2) [43]:

$$x_n = Ax_{n-1} + Bx_{n-2}. \tag{2}$$

The resulting Fibonacci numbers 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, …, have been the subject of continuing research, especially by the Fibonacci Association [40], publisher of the *Fibonacci Quarterly* since 1963, and its ramifications and applications are still growing in various areas of research, such as business and economics [26–28], biology [29–31], archeology [32–34], and mathematics [35–37].

In stock markets, Fibonacci numbers are commonly used in technical analysis with the knowledge of Elliott wave analysis to determine potential support, resistance, and stock price objectives [17]. The Elliott wave principle was introduced in the late 1920s by Ralph Nelson Elliott. He discovered that stock markets did not behave in a chaotic manner, but that markets move in repetitive cycles, which reflect the actions and emotions of humans, caused by exterior influences or mass psychology. Elliott contended, that the ebb and flow of mass psychology always revealed itself in the same repetitive patterns, which he was able to subdivide into so called "waves" [15].

The impulsive waves that Elliott discovered are built in basic structures (see Fig. 1 [45]) and always show five waves in stock markets. There are always two modes of the wave development, motive and corrective, to a complete pattern (one cycle composed of an up-and-down trend, namely, a bull market and a bear market such as Figs. 2 and 3 [45]). It was discovered that most patterns are composed of specific waves, matching Fibonacci numbers. These recurring patterns can be employed in forecasting stock market prices, based on the numbers [15,16].

Elliott has concluded that the Fibonacci summation series is the basis of the wave principle. Numbers from the Fibonacci sequence surface repeatedly in Elliott wave structures, including motive waves (1, 3, 5), a single full cycle (5 up, 3 down = 8 waves), and the completed motive (89 waves) and corrective (55 waves) patterns [16,38].

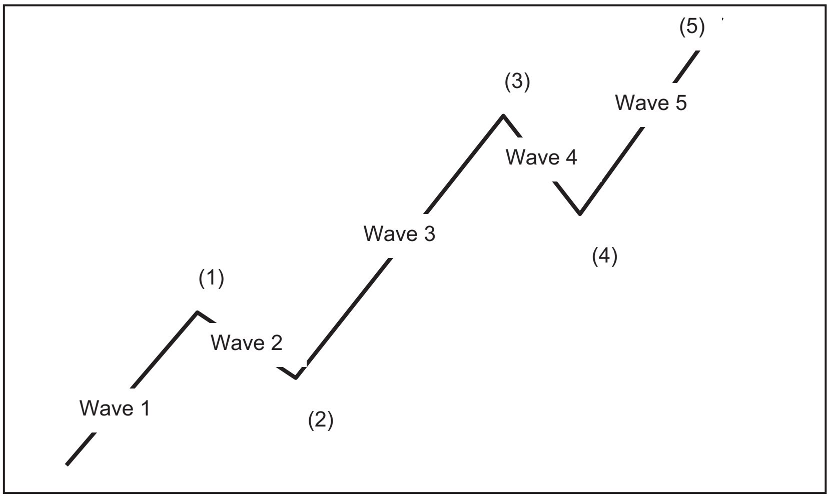

Fig. 1. Basic pattern of Elliott wave principle.

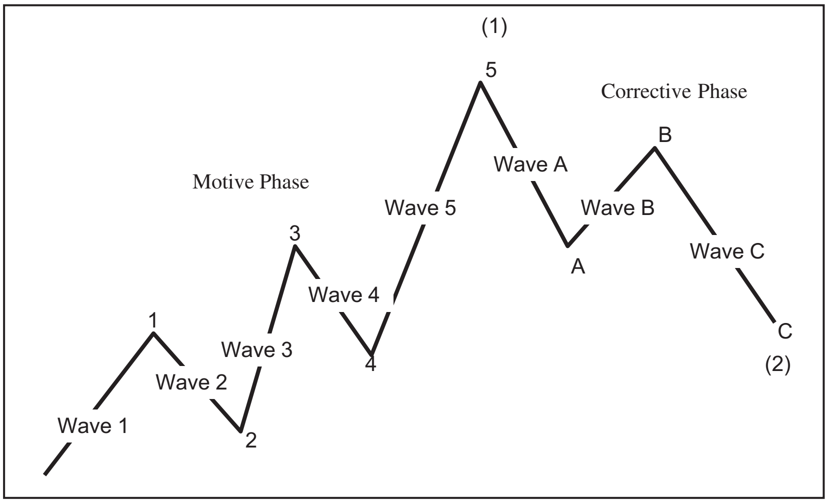

Fig. 2. Modes of the wave development: motive and corrective.

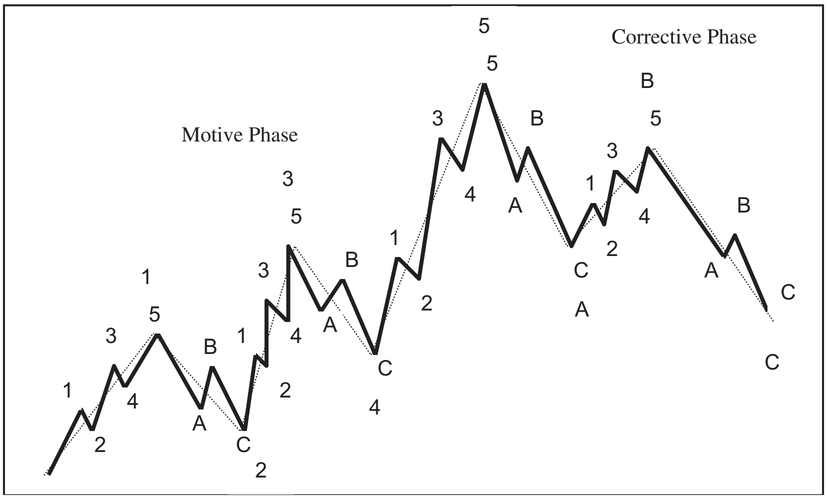

Fig. 3. Fractal pattern of Elliott wave principle.

3. Proposed model

In this section, the researchers propose a new fuzzy time-series model based on the Fibonacci sequence (see Fig. 4) and introduce the algorithm of the proposed model.

3.1. Research model

From the reviewed literature, we have found that the issues, such as determining a reasonable universe of discourse, discovering proper recurrence information, inspecting the data distribution within linguistic values and applying stock analysis theory in fuzzy time-series models, can be further discussed. To deal with these problems at one time, we propose a new fuzzy time-series model based on the Fibonacci sequence. There are three refined processes factored into our model

(1) The use of a fluctuation-weighted method to represent the patterns of stock price fluctuations in history.
(2) The use of the spread center of each linguistic value as a defuzzified value.
(3) The application of the Fibonacci sequence, $F_n = F_{n-1} + F_{n-2}$, to the forecasting process.

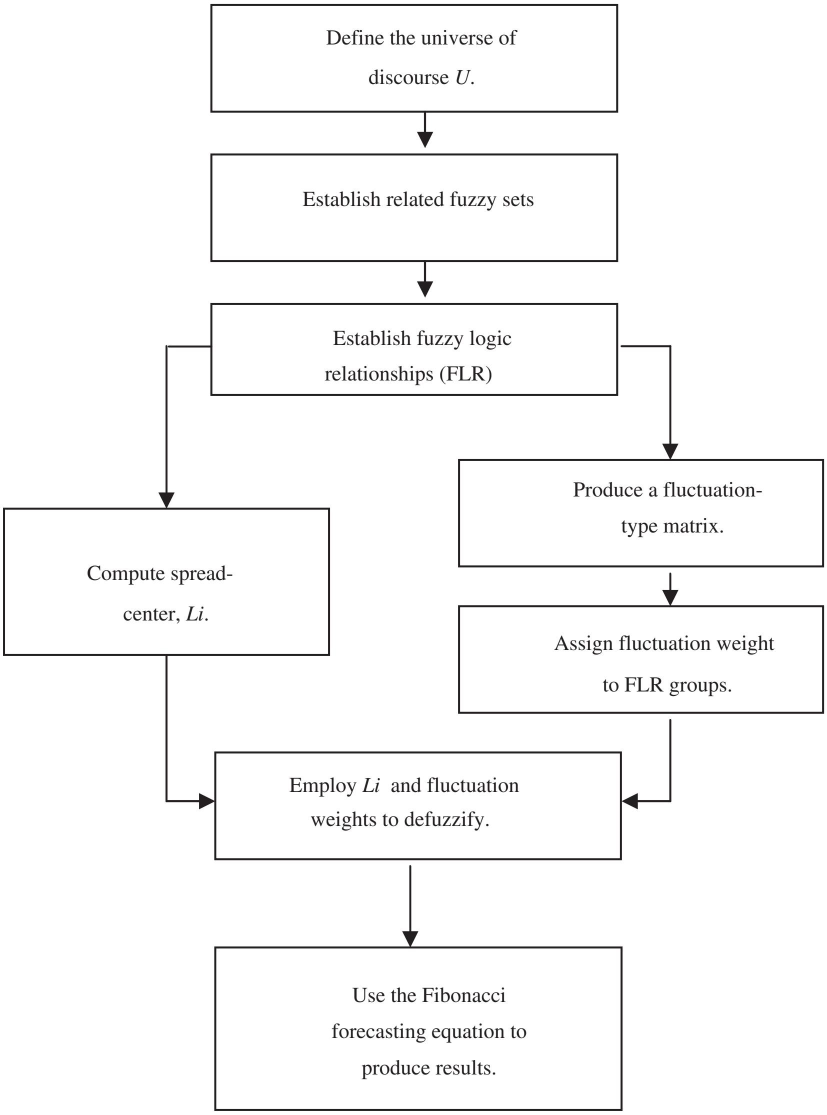

Fig. 4. Research process of the proposed model.

The first comes from the belief that a properly weighted method for mining the fuzzy logic relationships (FLRs, which are defined in step 3 of the proposed algorithm) in time-series will improve the forecasting accuracy of conventional models. Traditionally, the weight assigned for each FLR is determined either based on chronological order or experts' knowledge. Since most FLRs occur again within a specific time period, the history of the occurrence frequency can be used to estimate probability of reoccurrence in the near future. Therefore, classifying FLRs and weighting their occurrence is a reasonable approach for making predictions [10,39].

Secondly, the midpoint of the linguistic interval is employed in "defuzzification" to produce forecasting results in several fuzzy time-series models [1,10,39]. Although this method is an easily computed and often adopted approach, it ignores the data distribution of the observations within each linguistic interval and will result in large forecasting errors, if the data distribution is not symmetrical. Therefore, applying the spread-center of observations, the mean of observations, as defuzzified values should reconcile the problem of data distribution in some cases. For example, if there are five observations, 11, 12, 12, 13, and 17, falling in a linguistic interval, I, which ranges from 10 to 20, the mean of observations is 13 and the midpoint of the linguistic interval is 15. In defuzzification process, the forecasting errors using the mean of observations for these five observations are 2, 1, 1, 0, and 4, respectively, and the average error is 0.8. However, the forecasting errors using the midpoint are 4, 3, 3, 2, and 2, respectively, and the average error is 2.8. Under this circumstance of asymmetric data distribution, using the mean of observations in defuzzification will produce less forecasting error.

Thirdly, applying stock analysis theory is another reasonable approach to improve forecasting accuracy, whenever the forecasted targets are financial materials. The Elliott wave principle utilizes the Fibonacci sequence to predict the *timing* of stock price fluctuations and has become one of the major stock technical analysis theories [15,16,40]. Therefore, applying this sequence to a fuzzy time-series model is a theoretically reasonable approach. By way of summarization, we crystallize these refined processes into our model and provide the algorithm in the next subsection.

3.2. The algorithm

In this subsection, we detail the algorithm of the proposed model by using a 5-year period of TSMC stock data, from 2000 to 2004 (see Fig. 5) and several numerical examples to explain each step.

Step 1: Reasonably define the universe of discourse and decide into how many intervals the universe will be partitioned. There are two parts in this step as follows:

(1) For example, if the minimum and maximum TSMC stock price in the training dataset from 2000/01/04 to 2002/12/31 (see Fig. 5) is 36 and 219, respectively, the universe of discourse to be defined as $U = [35, 220]$. By using the fuzzy method, the low bound can be expanded by 1 smaller than 36, making *Low* 35, and up bound can be expanded by 1 larger than 219, making *Up* 220. As a result, the defined universe of discourse, $U = [35, 220]$, can cover every occurring stock price in the training dataset.
(2) Initially, partition the universe of discourse into seven linguistic intervals [23], and calculate the average datum falling within each. If the occurrence of the observations in a specific linguistic interval, I_k (where $k = 1, 2, 3, \ldots, n$; the initial value of n is 7) is greater than the average amount of all linguistic intervals, then the linguistic interval should be split a second time, into two more linguistic intervals. We provide a simple algorithm to describe this partitioning process as follows [39].

While all linguistic intervals have not been examined do
  If the occurrence of observations in $I_k >$ average occurrence of n linguistic intervals, then
(1) divide I_k into two equal length of second-divided intervals.
(2) update n and average occurrence.

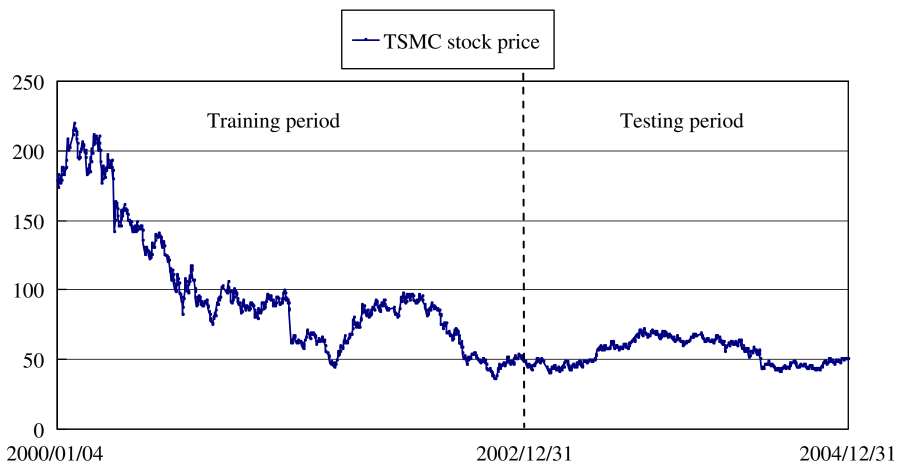

Fig. 5. TSMC stock price from 2000/01/04 to 2004/12/31.

Else examine next linguistic interval.
The number of partitioning intervals, n, is determined.

For example, the initial seven partitioning linguistic intervals for the universe of TSMC stock price are demonstrated in Table 1.

Step 2: Establish a related fuzzy set (linguistic value) for each observation in the training dataset. In this step, the fuzzy sets, $L_1, L_2, \ldots, L_k$, for the universe of discourse are defined by

$$\begin{aligned} L_1 &= a_{11}/u_1 + a_{12}/u_2 + \cdots + a_{1m}/u_m \\ L_2 &= a_{21}/u_1 + a_{22}/u_2 + \cdots + a_{2m}/u_m \\ &\vdots \\ L_k &= a_{k1}/u_1 + a_{k2}/u_2 + \cdots + a_{km}/u_m. \end{aligned} \tag{3}$$

For example, in this paper, the seven fuzzy linguistic values $L_1 =$ (very low price), $L_2 =$ (low price), $L_3 =$ (little low price), $L_4 =$ (normal price), $L_5 =$ (little high price), $L_6 =$ (high price) and $L_7 =$ (very high price) are applied [3]

$$\begin{aligned} L_1 &= 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + 0/u_5 + 0/u_6 + 0/u_7 \\ L_2 &= 0.5/u_1 + 1/u_2 + 0.5/u_3 + 0/u_4 + 0/u_5 + 0/u_6 + 0/u_7 \\ L_3 &= 0/u_1 + 0.5/u_2 + 1/u_3 + 0.5/u_4 + 0/u_5 + 0/u_6 + 0/u_7 \\ L_4 &= 0/u_1 + 0/u_2 + 0.5/u_3 + 1/u_4 + 0.5/u_5 + 0/u_6 + 0/u_7 \\ L_5 &= 0/u_1 + 0/u_2 + 0/u_3 + 0.5/u_4 + 1/u_5 + 0.5/u_6 + 0/u_7 \\ L_6 &= 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0.5/u_5 + 1/u_6 + 0.5/u_7 \\ L_7 &= 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0/u_5 + 0.5/u_6 + 1/u_7. \end{aligned} \tag{4}$$

In Eq. (3), the value of a_{ij} indicates the grade of membership of u_j in fuzzy set L_i, where $a_{ij} \in [0, 1]$, $1 \leqslant i \leqslant k$ and $1 \leqslant j \leqslant m$. Ascertain the degree of each stock price belonging to each L_i ($i = 1, \ldots, m$). If the maximum membership of the stock price is under L_k, then the fuzzified stock price is labeled as L_k. Each observation in the training dataset can be classified by the seven partitioned intervals (see Table 1) and labeled with a specific linguistic value. Table 2 demonstrates the assignments of linguistic value for eight periods of stock prices based on the classifications from Table 1.

In fuzzy time-series, the fuzzy logical relationships are generated based on the fuzzified stock price, L_k.

Step 3: Establish FLRs for linguistic time-series. One fuzzy logical relationship (FLR) is composed of two consecutive linguistic values. For example, the FLR ($L_i \rightarrow L_j$) is established by $L_i(t-1)$ and $L_j(t)$. Table 3 demonstrates the FLR, establishing processes of the linguistic time-series based on Table 2.

Step 4: Establish FLR groups and produce a fluctuation-type matrix. The FLRs with the same left hand side (LHS) linguistic value can be grouped into one FLR group. All FLR groups will construct a fluctuation-type matrix. Table 4 shows the fluctuation-type matrix produced by the FLRs from Table 3. Each row of the matrix represents one FLR group and each cell represents the occurrence frequency of each FLR.

Table 1
Seven partitioning intervals of TSMC

Partitioning intervals	Range
I_1	[35, 61]
I_2	[61, 88]
I_3	[88, 114]
I_4	[114, 141]
I_5	[141, 167]
I_6	[167, 194]
I_7	[194, 220]

Table 2
Assign a related linguistic value to each stock price

Time	Stock price	Linguistic value
$t = 1$	36	L_1
$t = 2$	37	L_1
$t = 3$	62	L_2
$t = 4$	50	L_1
$t = 5$	66	L_2
$t = 6$	90	L_3
$t = 7$	100	L_3
$t = 8$	120	L_4

Table 3
FLR table

$L_1(t = 1) \rightarrow L_1(t = 2)$
$L_1(t = 2) \rightarrow L_2(t = 3)$
$L_2(t = 3) \rightarrow L_1(t = 4)$
$L_1(t = 4) \rightarrow L_2(t = 5)$
$L_2(t = 5) \rightarrow L_3(t = 6)$
$L_3(t = 6) \rightarrow L_3(t = 7)$
$L_3(t = 7) \rightarrow L_4(t = 8)$

Table 4
An example of occurrence frequency of each FLR

$P(t)$	$P(t+1)$						
	L_1	L_2	L_3	L_4	L_5	L_6	L_7
L_1	1	2	0	0	0	0	0
L_2	1	0	1	0	0	0	0
L_3	0	0	1	1	0	0	0
L_4	0	0	0	0	0	0	0
L_5	0	0	0	0	0	0	0
L_6	0	0	0	0	0	0	0
L_7	0	0	0	0	0	0	0

Step 5: Assign fluctuation weight for each FLR group [39]. Each FLR within the same FLR group should be assigned a weight. For example, in Table 4, the FLR group of L_1 is $L_1 \rightarrow L_1, L_2$. The FLR of $L_1 \rightarrow L_2$ occurs once and the weight is assigned 1. However, The FLR of $L_1 \rightarrow L_2$ occurs twice. Therefore, the first FLR is assigned 1 and the second FLR is assigned 2. In the weighted method, the FLR weight is determined by its order of occurrence. Here, we provide an equation to compute the fluctuation weight assigned for each FLR, which is defined in (5) (where $W_{L_i \rightarrow L_j}$ represents the assigned weight for the FLR, $L_i \rightarrow L_j$; $f(L_i \rightarrow L_j)$ is the occurrence frequency of $L_i \rightarrow L_j$; and k is the occurrence order of $L_i \rightarrow L_j$).

$$W_{L_i \rightarrow L_j} = \sum_{k=0}^{f(L_i \rightarrow L_j)} k. \tag{5}$$

For example, Table 5 demonstrates the fluctuation-weighted matrix converted from Table 4.

Table 5
An example of fluctuation-weighted matrix

$P(t)$	$P(t+1)$							
	L_1	L_2	L_3	L_4	L_5	L_6	L_7	$\sum_{k=1}^{i} W_k$
L_1	1	3	0	0	0	0	0	4
L_2	1	0	1	0	0	0	0	2
L_3	0	0	1	1	0	0	0	2
L_4	0	0	0	0	0	0	0	0
L_5	0	0	0	0	0	0	0	0
L_6	0	0	0	0	0	0	0	0
L_7	0	0	0	0	0	0	0	0

Table 6
An example of defuzzification process for generating initial forecasts

Time	Stock price	Linguistic value	Defuzzify: $L_{df}(t) \times W_n(t)$ 1	Forecast $(t+1)$
$t=1$	36	L_1	$[41,64,95,120,0,0,0] \times [1/4,3/4,0,0,0,0,0]_T = 58.25$	
$t=2$	37	L_1	$[41,64,95,120,0,0,0] \times [1/4,3/4,0,0,0,0,0]_T = 58.25$	58.25
$t=3$	62	L_2	$[41,64,95,120,0,0,0] \times [1/2,0,1/2,0,0,0,0]_T = 68$	58.25
$t=4$	50	L_1	$[41,64,95,120,0,0,0] \times [1/4,3/4,0,0,0,0,0]_T = 58.25$	68
$t=5$	66	L_2	$[41,64,95,120,0,0,0] \times [1/2,0,1/2,0,0,0,0]_T = 68$	58.25
$t=6$	90	L_3	$[41,64,95,120,0,0,0] \times [0,0,1/2,1/2,0,0,0]_T = 107.5$	68
$t=7$	100	L_3	$[41,64,95,120,0,0,0] \times [0,0,1/2,1/2,0,0,0]_T = 107.5$	107.5
$t=8$	120	L_4	No rules	107.5

The sum of the weight of each FLR should be standardized to obtain a fluctuation-weighted matrix, $W_n(t)$. Eq. (6) defines the standardized weight matrix, $W_n(t)$:

$$W_n(t) = \left[W'_1, W'_2, \ldots, W'_i\right] = \left[\frac{W_1}{\sum_{k=1}^{i} W_k}, \frac{W_2}{\sum_{k=1}^{i} W_k}, \ldots, \frac{W_i}{\sum_{k=1}^{i} W_k}\right]. \tag{6}$$

For example, the standardized weights for L_1 in Table 5 are specified as follows: $W_1 = 1/4$, $W_2 = 3/4$, $W_3 = 0$, $W_4 = 0$, $W_5 = 0$, $W_6 = 0$ and $W_7 = 0$.

Step 6: Compute each spread center of linguistic value. The spread center of the observations, v_i, for the linguistic value L_i, is defined as (u_{ik} is the membership function of linguistic value L_i; x_k is the observation value):

$$v_i = \frac{\sum_{k=1}^{n} u_{ik} x_k}{\sum_{k=1} u_{ik}}. \tag{7}$$

The linguistic spread center matrix, $L_{df}(t)$, can be generated by each spread-center of linguistic value. For example, based on Table 2, only three observations (36, 37, and 50) are classified as L_1. Therefore, the spread-center of L_1, v_1, is $(36+37+50)/3$. In this way, we can produce the linguistic spread-center matrix [41, 64, 95, 120, 0, 0, 0].

Step 7: Defuzzify. From steps 5 and 6, the standardized weight matrix and linguistic spread-center matrix can both be obtained. In this step, the two matrixes are used to generate initial forecasts. This process is called ‘‘defuzzification’’ (defined in Eq. (8)). For example, based on Tables 2 and 5, the initial forecasts are produced and demonstrated in Table 6.

$$\text{Forecast}(t+1) = L_{df}(t) W_n(t). \tag{8}$$

Table 7
The forecasting performance of the proposed model (TSMC)

Dataset	Time period	Days	(α, β)	RMSE
Testing	2003/01/04–2003/06/30	117	(0.1, −0.1)	1.27
	2003/07/01–2003/12/31	130		1.33
	2004/01/02–2004/06/30	122		1.70
	2004/07/01–2004/12/31	128		0.86
Total	2003/01/04–2004/12/31	497		1.32

Step 8: Use the Fibonacci forecasting equation (defined in Eq. (9)) to produce conclusive forecasting results. By using the two linear parameters, α and β, as adjustments, the proposed model can adapt the forecasting results with the possibility of minimal error in the training dataset, so as to make our forecasts more reliable:

$$\begin{aligned}\text{Fibonacci\_forecast}(t+1) = P(t) + \alpha(\text{Forecast}(t+1) - P(t)) \\ + \beta(\text{Forecast}(t) - P(t-1)).\end{aligned} \tag{9}$$

In this step, α and β are adapted one by one from -1 and 1 but not including 0. The parameters are determined when the forecasting performance reach best in the training dataset. Then, we can use the adapted parameters to forecast for the target testing dataset. For example, the suggested parameters, α and β, for the example TSMC dataset are 0.1 and -0.1 (see Table 7). In practical applications, the values of the parameter set may vary in different training datasets because of the variations of data distributions. Based on this forecasting equation, the forecasting results from our model will fit recent stock price fluctuations more closely and, therefore, are more reliable to forecast the future prices.

4. Model verifications and comparisons

In this section, we employ two types of stock data, TSMC (Taiwan Semiconductor Manufacturing Company) stock price and TAIEX (Taiwan Stock Exchange Capitalization Weighted Stock Index) stock index [41] as experimental datasets to verify the forecasting performance of the proposed model.

4.1. Forecasting for TSMC stock price

A 5-year period of TSMC stock data (see Fig. 5) is employed to initially test the proposed model. The 3-year period of the TSMC dataset, from 2000/1/4 to 2002/12/31, is used for training, and the rest, from 2003/1/4 to 2004/12/31, for testing. To evaluate forecasting performance, RMSE (root mean square error, defined in Eq. (10)) is employed as a performance indicator [10,24]:

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{n} |\text{actual}(t) - \text{forecast}(t)|^2}{n}}. \tag{10}$$

Using seven linguistic values, we produce forecasting results for the TSMC stock prices within the testing periods (illustrated in Figs. 6–9) and a performance table with these forecasting values (see Table 7).

To verify forecasting performance, Chen's [3] and Yu's [10] models are used as comparison models. With the duplicated algorithms of these two models from literatures [3,10], we produce a forecasting performance comparison table (demonstrated in Table 8). It is apparent that the proposed model bears the smallest RMSE among three models (the RMSE of the proposed model is 1.32, Chen's model is 12.77 and Yu's model is 6.72). Based on this excellent performance, the further experimentation with a large-scale financial dataset is warranted.

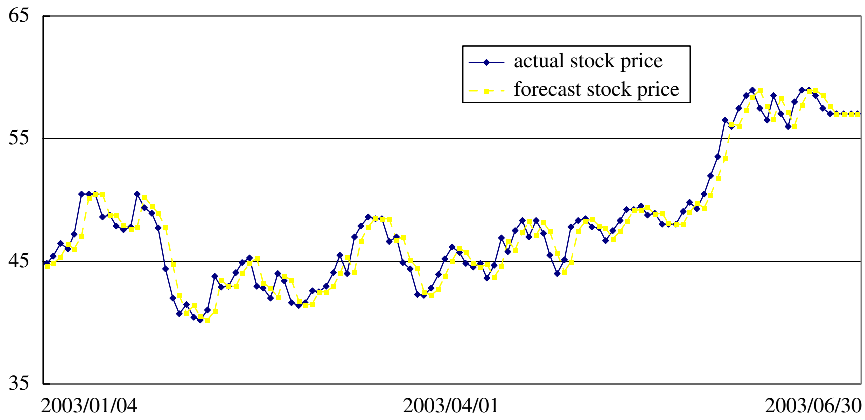

Fig. 6. The forecasting results (2003/01/04–2003/06/30).

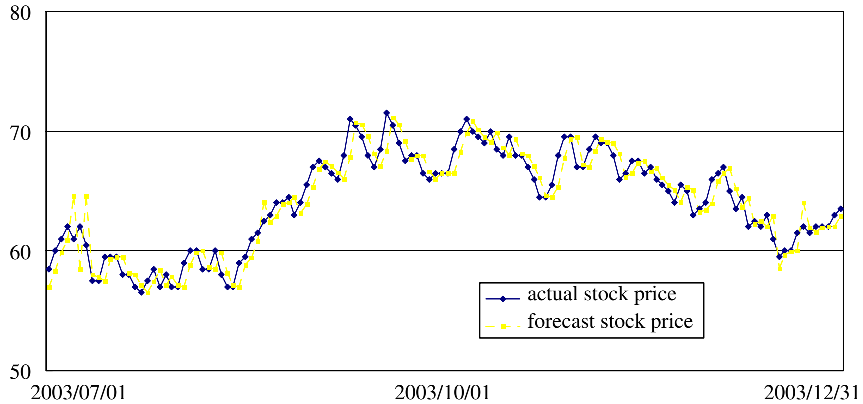

Fig. 7. The forecasting results (2003/07/01–2003/12/31).

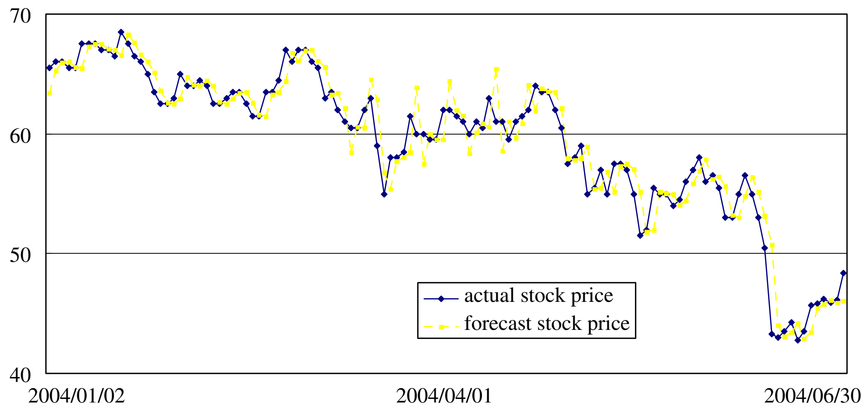

Fig. 8. The forecasting results (2004/01/02–2004/06/30).

4.2. Forecasting for TAIEX stock index

To verify the improvement of the proposed model, the further experimentation, using another large scale of financial dataset, is implemented. In this experiment, a 13-year period of TAIEX data, from 1991 to 2003, is

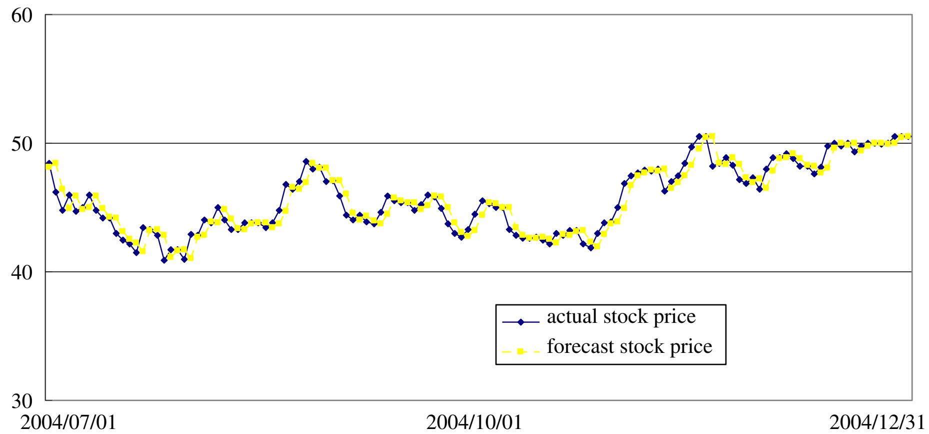

Fig. 9. The forecasting results (2004/07/01–2004/12/31).

Table 8
The forecasting performance comparisons

Models	Data set	Time period	Days	(α, β)	RMSE
Chen's model [3]	Testing	2003/1/4–2004/12/31	497	None	12.77
Yu's model [10]	Testing	2003/1/4–2004/12/31	494	None	6.72
Proposed model	Testing	2003/1/4–2004/12/31	497	(0.1, −0.1)	1.32

selected. Previous 10-month of each year, January to October, is used for training and the rest, November to December, for testing [10,24].

Using seven linguistic values, we generate 13 forecasting performances for all testing periods. To examine whether the proposed method outperforms the latest and conventional fuzzy time-series models, we employ Chen's [3] and Huarng's [24] models as comparison models. Using the performance data from Huarng's research [24], we produce a performance comparison table (see Table 9) and figure (see Fig. 10). Table 9 shows that the proposed model bears the smallest RMSE in every testing period.

5. Findings and discussions

From verification section, we can see that the proposed model outperforms Chen's [3], and Huarng's [24] models in all testing periods using the TAIEX dataset. Although the proposed model employs a common stock analysis technique, the Fibonacci sequence, in forecasting processes, it surpasses the latest model using neural networks [24]. In this empirical analysis, there are two findings provided as follows:

(1) The weighted methods for mining FLRs provides a more reasonable description of past stock price fluctuation patterns [10,39]. Our model and Huarng's model employ different weighted methods to assign a weight to each FLR. However, Chen's model gives each FLR the same weight [3]. From Table 9, we can see that Chen' model performs the worst.
(2) The Fibonacci forecasting equation makes forecasts match recent price fluctuations closely and results in better forecasting performance. From Eq. (9), the conclusive forecasting value is composed of the present stock price, $P(t)$, and previous two periods of modified forecasting errors, Forecast$(t+1) - P(t)$ and $Forecast(t) - P(t-1)$. Table 9 shows that our model surpasses Chen's and Huarng's in all testing periods of the TAIEX dataset. Therefore, we argue that the Fibonacci forecasting equation can modify our forecasts to fit recent price fluctuations more closely than those generated by conventional models.

Table 9
The performance comparisons of different models (TAIEX)

Models	Year													
	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	Average
Chen's model [3]	71	61	107	107	65	64	154	134	120	176	148	101	74	106
Huarng's model [24]	54	54	107	79	74	73	141	121	109	152	130	84	56	95
Proposed model	44[a]	43[a]	98[a]	76[a]	56[a]	50[a]	135[a]	114[a]	103[a]	128[a]	116[a]	66[a]	55[a]	83

[a]Performs best among three models.

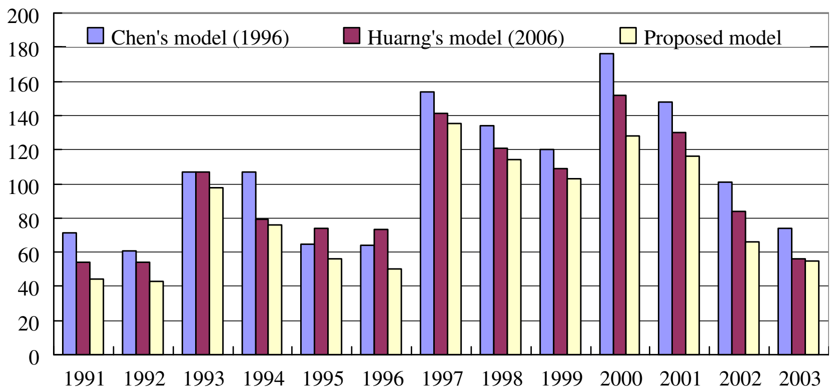

Fig. 10. The performance comparisons of different models (TAIEX).

Additionally, in practical applications, the proposed model can support investors to make investment decisions in stock markets. Investors can make their decisions, buying or selling stocks, based on a daily forecast generated from this model. If the forecasting closing price is going upside, the investors should buy stocks at the opening price and sell stocks at the closing price to make profits. On the contrary, the investors should sell stocks at the opening price and buy back the sold stocks if the forecasting price is going downside.

6. Conclusions and future research

In this paper, a new fuzzy time series, which incorporates the Fibonacci sequence and weighted method in forecasting processes to improve forecasting accuracy, has been proposed. Based on the verification results, we conclude that the research goal has been reached. However, there is still room for testing and improving the hypothesis of this model as follows: (1) employing other stock and financial materials 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 the profit making and (3) reconsidering the factors affecting the behavior of the stock markets, such as trading volume, news and financial reports which might impact stock price in the future.

References

[1] Q. Song, B.S. Chissom, Forecasting enrollments with fuzzy time-series. Part I, Fuzzy Sets Syst. 54 (1993) 1–10.
[2] Q. Song, B.S. Chissom, Forecasting enrollments with fuzzy time-series. Part II, Fuzzy Sets Syst. 62 (1994) 1–8.
[3] S.M. Chen, Forecasting enrollments based on fuzzy time-series, Fuzzy Sets Syst. 81 (1996) 311–319.
[4] S.M. Chen, Forecasting enrollments based on high-order fuzzy time series, Cybern. Syst. 33 (2002) 1–16.
[5] S.M. Chen, C.C. Hsu, A new method to forecast enrollments using fuzzy time series, Appl. Sci. Eng. 2 (2004) 234–244.
[6] S.M. Chen, N.Y. Chung, Forecasting enrollments using high-order fuzzy time series and genetic algorithms, Int. J. Intell. Syst. 21 (2006) 485–501.

[7] J.R. Hwang, S.M. Chen, C.H. Lee, Handling forecasting problems using fuzzy time series, Fuzzy Sets Syst. 100 (1998) 217–228.
[8] K.H. Huarng, Heuristic models of fuzzy time-series for forecasting, Fuzzy Sets Syst. 123 (2001) 137–154.
[9] R.W. Faff, R.D. Brooks, Y.K. Ho, New evidence on the impact of financial leverage on beta risk: a time-series approach, North Am. J. Econ. Finance 13 (2002) 1–20.
[10] H.K. Yu, Weighted fuzzy time-series models for TAIEX forecasting, Physica A 349 (2004) 609–624.
[11] C.H. Leon Lee, A. Liu, W.S. Chen, Pattern discovery of fuzzy time series for financial prediction, IEEE Trans. Knowl. Data Eng. 18 (2006) 613–625.
[12] K.H. Huarng, H.K. Yu, A type-2 fuzzy time-series model for stock index forecasting, Physica A 353 (2005) 445–462.
[13] H. Dourra, S. Pepe, Investment using technical analysis and fuzzy logic, Fuzzy Sets Syst. 127 (2002) 221–240.
[14] Y.F. Wang, Predicting stock price using fuzzy grey prediction system, Experts Syst. Appl. 22 (2002) 33–39.
[15] K. Jordan, An introduction to the Elliott wave principle, Alchemist 40 (2004) 12–14.
[16] J.C. Charles, A.J. Frost, R. Robert Jr., Elliott Wave Principle: Key to Market Behavior, Wiley, New York, 2003.
[17] R. Fischer, The New Fibonacci Trader: Tools and Strategies for Trading Success, Wiley, New York, 2001.
[18] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning III, Inf. Sci. 9 (1976) 43–80.
[19] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning I, Inf. Sci. 8 (1975) 199–249.
[20] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning II, Inf. Sci. 8 (1975) 301–357.
[21] K.H. Huarng, Effective lengths of intervals to improve forecasting in fuzzy time-series, Fuzzy Sets Syst. 123 (2001) 387–394.
[22] G.A. Miller, The magical number seven, plus or minus two: some limits on our capacity of processing information, Psychol. Rev. 63 (1956) 81–97.
[23] H.K. Yu, A refined fuzzy time-series model for forecasting, Physica A 346 (2005) 657–681.
[24] K.H. Huarng, T.H.K. Yu, The application of neural networks to forecast fuzzy time series, Physica A 336 (2006) 481–491.
[25] V.E. Hoggatt Jr., Fibonacci and Lucas Numbers, The Fibonacci Association: Houghton Mifflin, 1969.
[26] B. Swancoat, E. Kasanjian, Forecasting market turns using static and dynamic cycles, Tech. Anal. Stocks Commodities September (1992) 74–78.
[27] R.R. Prechter, The Wave Principle of Human Social Behavior and the New Science of Socionomics, New Classics Library, Georgia, 1999.
[28] R. Fisher, Fibonacci Applications and Strategies for Traders, Wiley, New York, 1993.
[29] D.W. Thompson, On Growth and Form, McMillan, New York, 1944.
[30] C. Sutton, Sunflower spirals obey laws of mathematics, New Sci. 18 (1992) 16.
[31] V.J. Roger, Mathematical Approach to Pattern and Form in Plant Growth, Wiley, New York, 1984.
[32] P. Hilton, J. Pedersen, Build Your Own Polyhedra, Addison-Wesley, Reading, MA, 1988, pp. 107–110, 113.
[33] R.E.M. Moore, Pattern formation in aggregations of entities of varied sizes and shapes as seen in Mosaics, Nature 209 (1996) 128–132.
[34] G.E. Runion, The Golden Section Dale Seymour, 2007.
[35] E.P.J. Miles, Generalized Fibonacci numbers and associated matrices, Am. Math. Monthly 67 (1960) 745–752.
[36] M. Wahl, A Mathematical Mystery Tour: Higher-Thinking Math Tasks, Zephyr Press, Tuscon, 1988.
[37] H.W. Gould, A history of the Fibonacci Q-matrix and a higher-dimensional problem, FQ 19 (1981) 250–259.
[38] R. Robert, R.N. Prechter Jr., Elliott's Masterworks: The Definitive Collection, New Classics Library, Georgia, 2005.
[39] 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.
[40] Fibonacci Association.
[41] TAIEX (Taiwan Stock Exchange Capitalization Weighted Stock Index) Stock Index.
[42] Elliott, 1938.
[43] Retrieved from the website: http://mathworld.wolfram.com
[44] Retrieved from the website: http://www.cs.rit.edu
[45] Retrieved from the website: http://www.goldenmuseum.com