

RESEARCH

Open Access



An integrated new threshold FCMs Markov chain based forecasting model for analyzing the power of stock trading trend

Kavitha Ganesan , Udhayakumar Annamalai and Nagarajan Deivanayagampillai

* Correspondence: kavithateam@gmail.com

Department of Mathematics,
Hindustan Institute of Technology &
Science, Padur, Chennai, Tamil Nadu
603103, India

Abstract

This paper explores the power of stock trading trend using an integrated New Threshold Fuzzy Cognitive Maps (NTFCMs) Markov chain model. This new model captures the positive as well as the negative jumps and predicts the trend for different stocks over 4 months which follow an uptrend, downtrend and a mixed trend. The mean absolute per cent error (MAPE) tolerance limits, the root mean square error (RMSE) tolerance limits are determined for various stock indices over a multi-timeframe period and observed for the existing methods lying within the defined limits. The results show for every 'n' number of predictions made, the predicted close value of the day's stock price was within tolerance limit with 0 % error and with 100% accuracy in predicting the future trend.

Keywords: Financial markets, Prediction intervals, Price forecasting, Comparative studies, Decision making, Fuzzy cognitive maps (FCMs), Markov chain

Introduction

Stock market is a place where uncertainty, vagueness, ambiguity, inconsistency and unpredictability prevail and persist. Price movements are observed daily in our day to day life. The behavior of stock prices appears to be rather odd. However, the price movements are not totally random in nature. Stock values are numerical values having uncertainty. Building a model for a problem involving uncertainty is no easy task. There are numerous forecasting tools and techniques available for predicting the stock market prices. Statistical models do not employ all available data for manipulation. The problem arises in handling unequally spaced, available historical high frequency data. Hence to address and resolve all the above challenges arising out of uncertainty and to achieve a better prediction accuracy, robust algorithms based on Markov chain and Fuzzy logic have been developed. Incidents of jumps are determined from the observations of the stock market data. This being an integrated model, it is more robust than the classical one and is well suited for analysis of high frequency data of observations.

A Markov chain is a stochastic process characterized by memoryless property. According to Markov chain concept, the probability distribution of the future state depends only on the present state and not on the sequence of happenings that precedes it. Markov chains are used in the area of finance to model a variety of different

phenomena including stock indices and market crashes. In 1974, Markov chains were used for modeling finance market behavior. Later, in 1989, they were used for modeling switches between periods of high volatility and low volatility of asset returns. A recent development is a Markov switching multi fractal model using Markov chains. In addition to these, Fuzzy logic analysis can be a very efficient tool in predicting stock market trend. Fuzzy models have the capability to analyze the datasets of linguistic values. Fuzzy set is a versatile tool for both linguistic and numerical modeling. For a randomly changing dynamic system, prediction of state of a Markov chain at a given point in the future is noted for uncertainty. In order to overcome this difficulty, the current model merges fuzzy logic tool with Markov concept for prediction and analysis.

The main purpose of this study is to integrate two major models, viz., the Markov chain model and the fuzzy model in an efficient way to obtain a solution for ambiguous situations like external factors affecting the stock market. A fuzzy approach has been developed, for analysis and study of the moving trends of the stock market leading to a generalized mathematical model called NTFCMs interrelated model and an integrated NTFCMs Markov chain model with internal threshold indicating the lower bound, upper bound and the exact threshold value, to enable palpable benefits to investors.

The model predicts the MAPE and RMSE tolerance limit for various stock indices. Comparisons are made with the other existing models measuring the effectiveness of the model for a multi-timeframe approach. The advantage of the model lies in the study of the positive, negative and mixed jumps for prediction analysis and for considering the most ambiguous situations in getting optimum solutions. Ambiguous situations are made possible for consideration since the historical price data are numerical values preprocessed and used for addressing the error tolerance effect and predicting the trend in the developed model. The interest here is to develop a low complexity, accurate prediction model which is better suited for real life situations. This study focuses on construction of an optimum solution using the interval partition in the Universe of discourse. Knowledge of the error tolerance limit, helps in taking reliable decisions predicting the trend of stock price index over varying timeframe. The estimates can be quite different depending on the frequency of the dataset.

This paper contains five sections. The First section provides the introduction. The Second section presents a review of literature. In the third and fourth sections, Markov chain and Fuzzy interrelated jump analysis prediction model, an integrated NTFCMs Markov chain model are defined respectively. Section 5 deals with applications of integrated NTFCMs Markov chain model to stock market moving trend analysis. Section 6 provides conclusions.

Review of literature

In recent years, fuzzy logic has been widely used as an efficient forecasting tool for solving problems of uncertainty prevailing in stock markets. Fuzzy set theory was first proposed by Zadeh in 1965. Kosko (1986) introduced the concept of fuzzy cognitive maps based on the structure of cognitive maps. Vasantha Kandasamy and Smarandache (2003) have analyzed the social problems in their book on fuzzy cognitive maps. Many models based on the concept of fuzzy time series were developed by researchers for forecasting the stock indices. Yu (2005), Yu and Huarng (2008), and Yu and Huarng

(2010) have proposed various fuzzy time series models for improvement of forecasting for Taiwan Futures Exchange (TAIEX). Chen and Chang (2010), Chen and Chen (2011), Chen et al. (2012), and Chen and Kao (2013) have developed various forecasting models for TAIEX based on fuzzy time series, fuzzy variation groups, fuzzy clustering, fuzzy rule interpolation technique, particle swarm optimization techniques and support vector machines. Wei et al. (2014) have proposed a hybrid ANFIS based on n -period moving average model to forecast TAIEX stock.

Huang et al. (2007) have proposed a multivariate heuristic model for fuzzy time-series forecasting. Wahab et al. (2013) have developed an interval type-2 fuzzy logic system model for measuring the index value of underground economy in Malaysia. Chen and Chen (2015) have proposed a hybrid fuzzy time series model based on granular computing for stock price forecasting. Hung (2011) has applied a combined fuzzy system and GARCH model to adaptively forecast stock market volatility. The concept of hybrid neuro fuzzy model for forecasting was developed by Madhoushi and Aliabadi (2012). Pritpal and Bhogeswar (2013) have proposed a high-order fuzzy-neuro expert system for time series forecasting. Hassan et al. (2012) has developed a hybrid fuzzy time series for forecasting. Li et al. (2010) has proposed a deterministic vector long-term forecasting for fuzzy time series. Crone et al. (2011) have made a 'Corrigendum to advances in forecasting with neural networks? Empirical evidence from the NN3 competition on time series prediction'.

The Hidden Markov Model (HMM) defined by Rabiner and Juang (1993) helps in tackling the situations involved in the consideration of the probability of the state caused by the previous state in forecasting. Rabiner (1993) provided a tutorial on Hidden Markov Model and selected applications in speech recognition. Rafiul et al. (2006), Hassan (2009), Hassan and Nath (2005), and Hassan et al. (2007) have proposed stock market for analyzing and predicting time series phenomena's using Hidden Markov Model (HMM), fuzzy model and extended their work with a fusion model of HMM, Artificial Neural Network (ANN) and Genetic Algorithm (GA) for forecasting. This study investigates the effectiveness of a hybrid approach. Cheng and Li (2012) combined fuzzy time series forecasting with a probabilistic smoothing hidden Markov model. Denkena and Nemeti (2013) have developed stock market related pricing mechanisms for the tool and mould manufacturing industry. Gharehchopogh et al. (2013) have provided a case study on linear regression approach for prediction of stock market trading volume. Makridakis and Hibon (2000) have discussed the M3-competition: results, conclusions and implications. Weron (2014) proposed electricity price forecasting: A review of the state-of-the-art with a look into the future. Park et al. (2016) have demonstrated a predicting market impact costs using nonparametric machine learning models. Iyer et al. (2015) have proposed a stock market prediction device using digital signal processing models. Xue et al. (2015) have introduced stock market trading rules based on the biclustering method. Guo et al. (2015) have proposed a stock market forecasting model combining a two-directional two-dimensional principal component analysis and a radial basis function neural network. Another effort towards the development of a new fuzzy time series forecasting model combined with ant colony optimization and auto-regression models for stock markets has been made by Cai et al. (2015). Gray et al. (2015) analyzed change, preferred states, and perceived resilience of social-ecological systems using fuzzy cognitive mapping as a participatory approach. Li

et al. (2016) proposed a group decision making model for integrating heterogeneous information. A new Fuzzy Cognitive Map learning Algorithm for Speech emotion recognition was developed by Zhang et al. (2017). Ceska et al. (2019) developed Shepherding Hordes of Markov Chains. Chao et al. (2019) have proposed behavior monitoring methods for trade-based money laundering integrating macro and micro prudential regulation: a case from china. Kou et al. (2019) has developed machine learning methods for systemic risk analysis in financial sectors. Zhang et al. (2019) has developed soft consensus cost models for group decision making and economic interpretations.

The existing literature review reveals employment of various tools for developing prediction models for different stock indices. These models have their own merits and demerits in solving problems of various domains. Expert knowledge is required for developing fuzzy systems. Recent advancements in research have provided new perspectives in the field of financial forecasting. According to the present study, a New Threshold Fuzzy Cognitive Maps (NTFCMs) Interrelated Model and an integrated NTFCMs Markov Chain Model, exhibiting an interval varying parameter dependent new threshold value are used in the very construction of the nodes of the dynamic system of the Fuzzy Cognitive Maps (FCMs). The first objective is to introduce a threshold parameter ' ϕ ' for model development. The second is to show the computational efficiency of the current model in providing the prediction performances within the tolerance range. The main objective is to develop an accurate forecasting model indicating lower bound, upper bound and the exact threshold values which can provide the best expected and the most optimal solution. For every ' n ' number of predictions made the predicted value is within tolerance limit with zero percent error and 100% accuracy. The results show the outcome of the new model, getting a forecasting accuracy per cent equivalent to other existing models. These features make the model well suited for financial market applications.

Markov chain and fuzzy interrelated jump analysis prediction model

Consider the states S_1, S_2, \dots, S_n such that S_i takes a value in the interval (a_i, a_{i+1}) for $1 \leq i \leq n$. So our Concepts or nodes C_1, C_2, \dots, C_n of the FCMs [5,14] do not take the usual form of adjective terminology expressing the nature of the concepts in a new linguistic way but they take the values at any point in the interval (a_i, a_{i+1}) . So, at this juncture we define nodes $S_i (i = 1, 2, \dots, n)$ which take values in the interval as Interval Nodes. This is the marked difference between the usual FCMs and the NTFCMs with entries from an interval.

These concepts in our new model take values in intervals, where these intervals are themselves varying states like Very Low, Minimum Low, Low, Moderate Low, Minimum High, High, Moderate High, Very High etc. which take their values from the stipulated intervals. So, now with this S_1, S_2, \dots, S_n as the concepts or nodes, a like of FCMs called as NTFCMs interrelated model is built. This new model caters to the need of the prefixed or pre-assumed values from an interval, where the pre-assumed values vary in this interval.

The states of the stock market can be taken according to expert decision and framed as the main co-ordinates for our dynamic system. For instance, for a dynamic system having six nodes that is $n = 6$; the states can be defined as S_1 - Very Low, S_2 - Low, S_3 - Moderate Low, S_4 - Moderate High, S_5 - High, S_6 - Very High. There may be need for

another expert to consider a dynamic system having 8 nodes say; S_1 - Very Low, S_2 - Minimum Low, S_3 - Low, S_4 -Moderate Low, S_5 -Minimum High , S_6 - High, S_7 -Moderate High, S_8 - Very High. The NTFCMs can take any number of nodes depending on the expert's wishes.

This method works on positive leaps or jumps as in the case of Markov Chains. When the leaps or jumps are in the anticlockwise direction, these jumps or leaps are dissolved as inconsistent or disobedient. Hence, this is not taken into consideration when a trend prediction is performed. This method is a judicious amalgam of both FCMs and Markov Chain models to enable the close value prediction of the actual trend in any problem. When a prediction of any trend in the problem is attempted, there cannot be any fixed value. Henceforth, we proceed to describe the dynamic system setup at the outset by using the concepts or attributes or nodes as intervals which, in turn, are very carefully defined describing a minute transition of one state to another indicating the jumps are analytically tractable.

The closing price for ' N ' number of days is used as input parameter for analyzing the stockmarket trend. The observations of the closing price are denoted by $c_1, c_2, \dots, c_t, \dots, c_N$. The difference variable defined by $c_{diff} = c_t - c_{t-1}$ is computed between every two consecutive t^{th} and $(t-1)^{th}$ day closing price stock value. The non-stationary values are transformed into stationary values. When the non-stationary closing price values undergo transformation, there will be only $(N-1)$ stationary values for the ' N ' days considered. Determine the range ' R ' for the stationary values. The range has an upper bound (UB) and the lower bound (LB). UB is given by maximum range $R_{max} = \max(c_{diff} \text{ values})$ and LB denotes minimum range $R_{min} = \min(c_{diff} \text{ values})$. Divide the range R into n equal number of sub intervals. Each co-ordinate (a_t, a_{t+1}) of the subinterval is used to define our states considered. Now find the midpoint $x_t, 1 \leq t \leq n$ of each sub interval. This range of values of the dynamic system, determines the observed frequencies. With a view to determine the expected frequencies of the range, we give deviations called varying parameters ' ϕ ' in the interval $[-\alpha, \alpha]$ to the midpoints of the defined range which will give the best impact on the dynamic system. A special type of NTFCMs described earlier in this section is used in the very construction of the states of the dynamic system of the FCMs. Midpoints x_i 's undergo a change of scale in the threshold value with respect to a small arbitrary parameter in the interval $[-\alpha, \alpha]$, denoted by ' ϕ ' and produce an output ' Y '. The output ' Y ' thus formed is a sequence of pattern comprising of Decreasing symbol ' D ' and Increasing symbol ' I ' depending upon whether the output is negative or positive of the dynamic system.

Positioning of each x_i at each stage is described as under:

- (i) Construction of an interval
- (ii) Obtaining the interval points
- (iii) Uniform jumps are only valid
- (iv) Clockwise jumps are valid
- (v) Anticlockwise jumps are disobedient jumps or danger jumps in this problem of moving stock market.

Table 1 States and the corresponding intervals of the jump analysis method

States	Corresponding Intervals
S_1	(a_1, a_2)
S_2	(a_2, a_3)
S_3	(a_3, a_4)
⋮	⋮
⋮	⋮
⋮	⋮
S_n	(a_n, a_{n+1})

The ‘ n ’ states and their corresponding intervals are defined in Table 1. Midpoint of the state S_t is defined by $x_t = \frac{a_t+a_{t+1}}{2}, 1 \leq t \leq n$. The collection of all midpoints put together forms a set, called the *MidX* set, given by $MidX = \{x_1, x_2, x_3, x_4, \dots, x_n\}$. The output ‘ Y ’ after undergoing a change with respect to the varying parameter ‘ ϕ ’ is the set denoted by $Y = \{y_1, y_2, y_3, y_4, \dots, y_n\}$ and is defined by $y_i = x_i \pm \phi, i = 1, 2, \dots, n$ where ‘ ϕ ’ is a small arbitrary parameter in the interval $[-\alpha, \alpha]$ depending on the problem on hand (α , a real number chosen after the analysis of the problem). The framework design and the technical details of each stage of the current study are described as shown in Fig. 1.

The working of the model is now discussed. The assumptions indicated here are due to the merging of the Markov chain model with FCMs. Let $S_1, S_2, \dots, S_r, S_{r+1}, \dots, S_{n-1}, S_n$ be the coordinates of the New FCM with internal threshold model constructed. An instantaneous state vector ‘ S ’ given by $S = (S_1, S_2, \dots, S_i, \dots, S_n)$ where $S_i \in \{0, 1\}$ as represented in (1) is considered as an initial state vector, denoting the ON-OFF positions of the states.

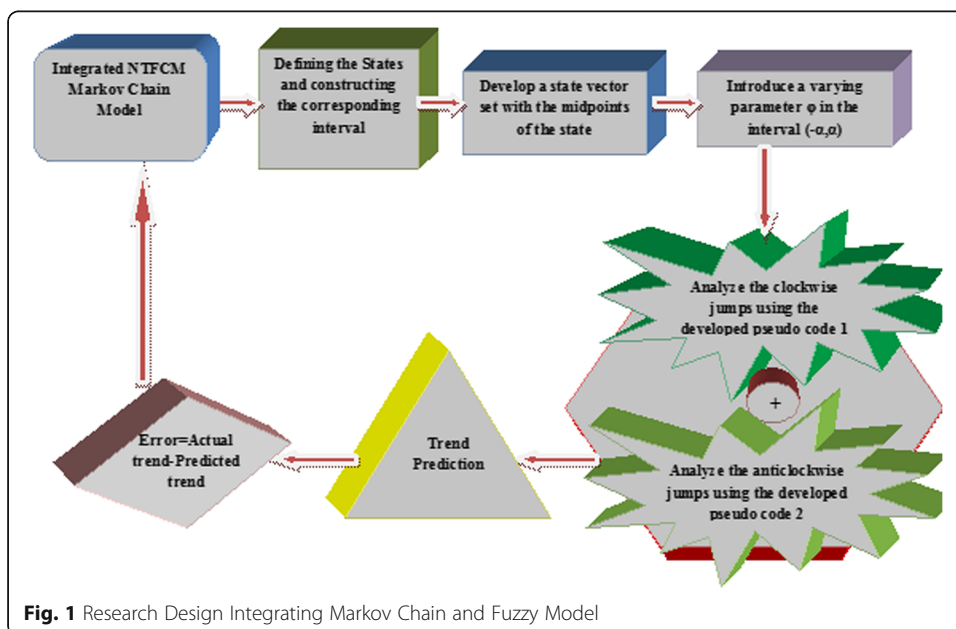


Fig. 1 Research Design Integrating Markov Chain and Fuzzy Model

$$S_i = \begin{cases} 1 & \text{if } S_i \text{ is in ON position} \\ 0 & \text{if } S_i \text{ is in OFF position} \end{cases} \text{ for } i = 1, 2, \dots, n \tag{1}$$

Consider a dynamic system. A Pseudo code has been developed for fitting the defined model. The level of the pseudo code determines the fitness of the optimization or the nature of the solution for the current method. The solution can be given ranks, the rank ranging from 5 to 0 as shown in Table 2.

Pseudo code 1:

```

for i = 1 : n - 1
for j = i + 1 : n
    = r
    = r - 1
    Sj - Si ≤ r - 2
    = r + 1
    ≥ r + 2
end
end
    
```

where “r” in the pseudo code is defined as:

$$r = \begin{cases} \frac{n}{2} + 1 & \text{when } n \text{ is even} \\ \frac{n+1}{2} & \text{when } n \text{ is odd} \end{cases}$$

Integrated NTFCMs Markov chain model

Under unusual cases, jumps or leaps can be on the other side. Such jumps are referred to as unnatural jumps. Concepts involving these jumps are ignored and avoided. Since we are using Markov chains, we call anticlockwise jumps as Disobedient jumps.

The negative jumps or leaps are always found in a dynamic system indicating a downward trend or a Bear State which shows the danger of falling. But, they are discarded or dissolved for the sake of convenience. But, one should consider the Disobedience Zone or Danger Zone to give a better model. The Danger Zone or the anticlockwise trend is the most important study because one can be cautious during that period.

Table 2 Range, rank and nature of the solution

Range		Rank	Nature of the Solution
Solution ' $S_j - S_i$ '	Reverse solution ' $S_i - S_j$ '		
r	$\leq r-2$	5	Most optimal solution
$r-1$	$r-1$	4	Optimal solution
$\leq r-1$	r	3	Just optimal solution
$r+1$	$r+1$	2	Ambiguous or rigged
$\geq r+2$	$\geq r+2$	1	Most ambiguous
$j=i$	$i=j$	0	Stationarity

In order to explain the Markov chain model, in addition to pseudo code 1 mentioned earlier, pseudo code 2 has been designed and it determines the nature of the solution.

Pseudo code 2:

```

for revi = n : 2
for revj = n - 1 : 1
    revsoln = {
        ≤ r - 2
        = r - 1
        = r
        = r + 1
        ≥ r + 2
    }

end
end

```

There can also be a combination of pseudo code 1 and pseudo code 2, considering the possibility of both positive and negative trends being followed.

The Markov chain denoted by 'mc' is defined using the conditions specified below:

```

If soln > revsoln,
mc = soln - revsoln
if mc < 0
    then assign mc = {
        -1
        -2
        -3
        ≤ -4
    }
if mc ≥ 0
    then assign mc = {
        0
        1
        2
        3
        ≥ 4
    }

```

The trend, rank and growth arrow corresponding to the mc values are indicated in Table 3.

Experimental result and discussion of integrated NTFCMs Markov chain model to the stock market moving trend problem

As a general strategy, it is best to trade with trends, indicating the need for caution when the general trend of the market is headed up. The caution relates to taking any positions that rely on the trend going in the opposite direction. A Bull Market Trend is

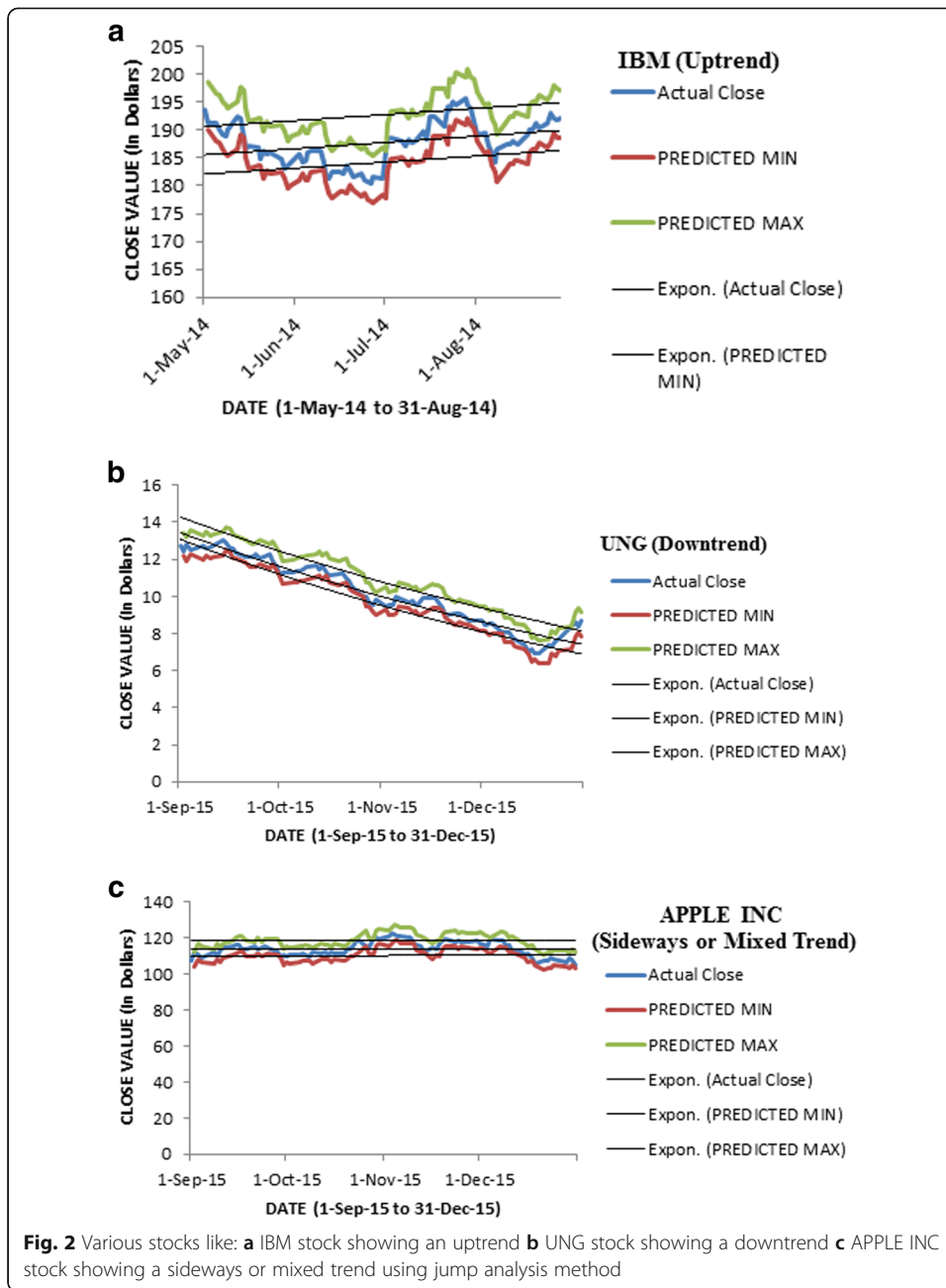
Table 3 Trend, rank and growth arrow for $mc < 0$ and $mc \geq 0$

mc value	Trend	Rank	Growth Arrow
≤ -4	Very less stable	1	↓
-3	Less stable	2	↓
-2	High stable	3	↓
-1	Very high stable	4	↓
0	Most stable	5	↑
1	Very high stable	4	↑
2	High stable	3	↑
3	Less stable	2	↑
≥ 4	Very less stable	1	↑

an Upward Trend associated with a long term climb with increasing investor confidence, and increased investment in anticipation of future price increases. It symbolizes charging ahead with excessive confidence. Price rise indicates a strong economy and a low level of unemployment. Market returns are good. A Bear Market Trend is a Downward Trend associated with long term decline with increasing investor fear level. It is a transition from high investor optimism to a widespread investor fear. It symbolizes lazy or cautious movers. Falling prices show a dull economy. In Recession, unemployment rates are high. Market's returns are bare. Analysis of the efficacy of the stock market moving trend for the methods developed, involves the consideration of three different stocks, one following an upward trend, a second showing a downward trend and the third a mixed trend to indicate this method works well under all circumstances - this is shown in Fig. 2. The results are compared with the existing trend and the accuracy of the model for various trends is found out. The current model is a multi-trend multi-timeframe model where the forecasting performance is tested for a time period of 1-day, 1-week, 2-weeks, 1-month, 2-months, 4 months ahead. There are two phases namely training and a testing phase. The data obtained for the stock indices consisted of close price in each day.

The stock prices for IBM, APPLE INC, UNG sectors at the close of every day for a period of one month are collected from the historical dataset available in www.finance.yahoo.com. In order to study the moving trend, the daily close value of the stock price of the IBM data set for a period of 1 month is collected. One day difference in close value is found out. The lower bound of the given range is found out. The upper bound of the given range is also found out by using the formula $h = \frac{b-a}{n}$. For the dataset considered, it is found out to be [-2.2, 1.88]. Similarly, the raw data of the various other sectors considered are also transformed into a corresponding range of values. The dataset considered is split into 6 subintervals, for instance, $n = 6$. Each subinterval is given a name as $S_1, S_2, S_3, S_4, S_5, S_6$. The Midpoints of the subintervals are found and it is given by $MidX = \{-1.86, -1.18, -0.5, 0.18, 0.86, 1.54\}$.

The midpoints of the subintervals are found out considering the possibility of accuracy in working each model. Instead of considering points from the full subinterval, midpoints are chosen in order to give a reliable working model. The midpoints thus obtained, form a state vector $\{X_1, X_2, X_3, X_4, X_5, X_6\}$. The state vector is made to undergo an arbitrary change with respect to a small arbitrary parameter ' ϕ ' in the interval [-2.2, 1.88], in order to determine the expected frequencies of the range



and produce an *OutputY*. Various outputs are obtained for varying 'φ'. When 'φ' is assumed a value 1 and added to the *MidX* state vector, the change in *MidX* state vector is noted. It is given by $OutputY = \{-0.86, -0.18, 0.5, 1.18, 1.86, 2.54\}$.

The change in the changed *MidX* i.e., *OutputY* is noted, for each corresponding values of *MidX*. When the change is positive, the trend arrow is indicated with a symbol denoted by '↑' and defined as increasing. If the change is negative, the trend arrow is indicated with a symbol denoted by '↓' and defined as decreasing. The hidden pattern from the *MidX* and *OutputY* can be determined. For instance the pattern vector is given by $\{\downarrow \downarrow \uparrow \uparrow \uparrow \uparrow\}$. We determine the connection matrix 'CM' for the *OutputY*. Table 4 gives the one day difference in close value for the IBM dataset considered:

Table 4 IBM stock close value dataset

<i>Closevalue</i>	<i>C_{diff}</i>	<i>Closevalue</i>	<i>C_{diff}</i>
\$77.91	–	\$79.07	0.51
\$77.39	– 0.52	\$77.40	–1.67
\$76.50	–0.89	\$77.28	–0.12
\$75.86	–0.64	\$77.95	0.67
\$77.45	1.59	\$77.33	–0.62
\$79.33	1.88	\$76.70	–0.63
\$79.51	0.18	\$77.73	1.03
\$79.15	–0.36	\$77.07	–0.66
\$79.95	0.8	\$77.90	0.83
\$78.56	–1.39	\$75.70	–2.2

The range is determined for Table 4 and is found to be [-2.2, 1.88]. The data set considered is split into 6 subintervals, say. Then for $N = 6$, $MidX = \{-1.86, -1.18, -0.5, 0.18, 0.86, 1.54\}$.

Also, the same study was done for $N = 8$, the $MidX$ obtained was $MidX = \{-1.945, -1.435, -0.925, -0.415, 0.095, 0.605, 1.115, 1.625\}$. The mid-points undergo a change with respect to a small arbitrary parameter ‘ ϕ ’ in the interval [- 2.2, 1.88], in order to determine the expected frequencies of the range and produce an *OutputY*. The *OutputY* in turn determines the hidden pattern of the dynamic system. The pattern depending on ϕ chosen is given by Y (i.e.,) the value of Y depends upon the positive or negative value of the output value y_i . For varying ϕ , we get varying patterns as shown in Table 5 from which conclusions are drawn. The Connection Matrix ‘CM’ Vasantha Kandasamy and Smarandache (2003) is formed for the y_i 's, $i = 1, 2, \dots, n$ using fuzzy cognitive mappings for the set considered. This pattern undergoes threshold and updation and the trend is analyzed.

Now for illustration, when $\phi = -0.25$

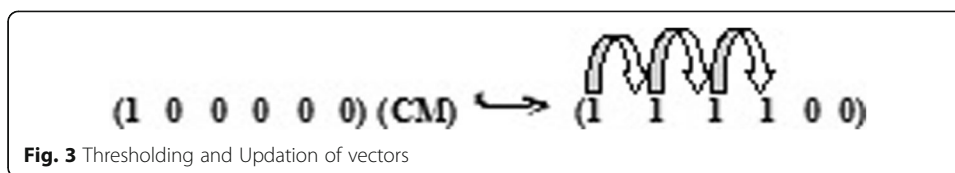
$$\begin{aligned}
 MidX &= \{-1.86, -1.18, -0.5, 0.18, 0.86, 1.54\} \\
 OutputY &= \{-2.11, -1.43, -0.75, -0.07, 0.61, 1.29\} \\
 \text{Let the Initial vector be } &(1 \ 0 \ 0 \ 0 \ 0 \ 0)
 \end{aligned}$$

$$CM = \begin{pmatrix} 0 & 1 & 1 & 1 & -1 & -1 \\ 1 & 0 & 1 & 1 & -1 & -1 \\ 1 & 1 & 0 & 1 & -1 & -1 \\ 1 & 1 & 1 & 0 & -1 & -1 \\ -1 & -1 & -1 & -1 & 0 & 1 \\ -1 & -1 & -1 & -1 & 1 & 0 \end{pmatrix}$$

↪ Thresholding and updation has been done for the vectors as shown in Fig. 3. For the dynamic system starting with S_1 switched ON, there are 3 jumps or leaps in the clockwise direction. After going through threshold and updation, S_2, S_3, S_4 are all ON from S_1 indicating the observed solution is in Rank 5 giving a Most Optimal Solution. Similarly, switching ON different S_i , many different solutions are observed as explained above in the Pseudo code of the current model and the trend of the stock market can

Table 5 Sample result for IBM stock dataset using jump analysis method

Sample Threshold Values of ' ϕ '	Output Pattern ' γ ' for N=6	Final Vector for (1 0 0 0 0 0)	Rank 5-0	Output Pattern ' γ ' for N=8	Final Vector for (1 0 0 0 0 0 0 0)	Rank
0.0078125	{↓ ↓ ↓ ↓ ↑ ↑ ↑ ↑}	(1 1 1 0 0 0)	4	{↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↑ ↑ ↑ ↑}	(1 1 1 1 0 0 0 0)	4
-0.0078125	{↓ ↓ ↓ ↓ ↓ ↑ ↑ ↑}	(1 1 1 0 0 0)	4	{↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↑ ↑ ↑ ↑}	(1 1 1 1 0 0 0 0)	4
0.1	{↓ ↓ ↓ ↓ ↓ ↑ ↑ ↑}	(1 1 1 0 0 0)	4	{↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↑ ↑ ↑ ↑}	(1 1 1 1 0 0 0 0)	4
-0.1	{↓ ↓ ↓ ↓ ↓ ↑ ↑ ↑}	(1 1 1 0 0 0)	4	{↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↑ ↑ ↑ ↑}	(1 1 1 1 0 0 0 0)	5
0.1875	{↓ ↓ ↓ ↓ ↓ ↑ ↑ ↑}	(1 1 1 0 0 0)	4	{↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↑ ↑ ↑ ↑}	(1 1 1 1 0 0 0 0)	4
-0.1875	{↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓}	(1 1 1 1 0 0)	5	{↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↑ ↑ ↑ ↑}	(1 1 1 1 1 0 0 0)	5
0.375	{↓ ↓ ↓ ↓ ↓ ↑ ↑ ↑}	(1 1 1 0 0 0)	4	{↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↑ ↑ ↑ ↑}	(1 1 1 1 0 0 0 0)	4
-0.375	{↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓}	(1 1 1 1 0 0)	5	{↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↑ ↑ ↑ ↑}	(1 1 1 1 1 0 0 0)	5
1.0625	{↓ ↓ ↓ ↑ ↑ ↑ ↑ ↑}	(1 1 0 0 0 0)	3	{↓ ↓ ↓ ↑ ↑ ↑ ↑ ↑ ↑ ↑}	(1 1 0 0 0 0 0 0)	3
-1.0625	{↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓}	(1 1 1 1 1 0)	2	{↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↑ ↑ ↑ ↑}	(1 1 1 1 1 0 0 0)	2
2	{↑ ↑ ↑ ↑ ↑ ↑ ↑ ↑}	(1 1 1 1 1 1)	1	{↑ ↑ ↑ ↑ ↑ ↑ ↑ ↓ ↓ ↓ ↓ ↓}	(1 1 1 1 1 1 1 1)	1
-2	{↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓}	(1 1 1 1 1 1)	1	{↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓}	(1 1 1 1 1 1 1 1)	1



be analysed. Using the Pseudo codes defined in “Markov chain and fuzzy interrelated jump analysis prediction model” and “Integrated NTFCMs Markov chain model” sections, the following results are obtained for the datasets considered. The predicted close price value c_t on t^{th} day is obtained using the formula given below.

Predicted value on the current t^{th} day = the observed value on the previous $(t - 1)^{th}$ day + $c_{diff}(t, t - 1)$.

$$\begin{aligned}
 Predictedmin_i &= Actualclose_i - R_{min} \\
 Predictedmax_i &= Actualclose_i - R_{max}
 \end{aligned}$$

The forecasting accuracy is measured in terms of the Root mean square error (RMSE), Mean absolute per cent error (MAPE). The following is the procedure to calculate the $RMSEmin$, $RMSEmax$, $MAPEmin$ and $MAPEmax$.

$$\begin{aligned}
 RMSE \min &= \sqrt{\frac{\sum_{i=1}^n (Predicted \min_i - Actualclose_i)^2}{n}} \\
 RMSE \max &= \sqrt{\frac{\sum_{i=1}^n (Predicted \max_i - Actualclose_i)^2}{n}} \\
 MAPE \min &= \frac{100}{n} \sum_{i=1}^n \frac{(Predicted \min_i - Actualclose_i)}{Actualclose_i} \\
 MAPE \max &= \frac{100}{n} \sum_{i=1}^n \frac{(Predicted \max_i - Actualclose_i)}{Actualclose_i}
 \end{aligned}$$

In the above defined formulas, ‘ n ’ denotes the total number of test data sequence; $Actualclose_i$ indicates the actual close value of the stock price on day i ; $Predictedmin_i$ and $Predictedmax_i$ indicates the predicted minimum close value, predicted maximum close value of the stock price on day i .

Table 6 Sample data for actual close value, predicted close value, tolerance, observed trend, existing trend and error for various stocks using Markov chain and fuzzy interrelated model

Stock Name	Actual close value	Predicted close value		Close value tolerance limit	Observed trend	Existing trend	Error	The Proposed model MAPE tolerance limit
		Predicted minimum close value	Predicted maximum close value					
IBM	\$186.35	\$184.7	\$190.22	(-1.65, 3.87)	Up trend	Up trend	Zero	
UNG	\$12.61	\$12.17	\$12.8	(-0.44, 0.19)	Down trend	Down trend	Zero	(-4.9,7.9)
APPLE INC	\$124.25	\$122.16	\$127.1	(- 2.09, 2.85)	Sideways or Mixed trend	Sideways or Mixed trend	Zero	

For every ‘*n*’ number of predictions made, the predicted maximum and the predicted minimum are taken to ensure it sticks on to the accurate one. When it is within tolerance limit, error is assumed to be zero. Table 6 gives the sample data for actual close price of the index on each day of the forecasting period, predicted close value, tolerance, observed trend, existing trend and error for various stocks considered. It is found that the observed trend and the existing trend obtained were same with Zero per cent Error and accuracy was found to be 100%.

Stock prices of firms in the IT sector have been used for demonstrating the authenticity of the proposed model: the daily stock price of Apple Computer Inc., International Business Machines Corporation (IBM) and Dell Inc., collected from www.finance.yahoo.com.

Table 7 shows the comparison of the current method with other existing methods for a time frame of 4 months. Figure 4 shows the Existing methods of Hassan, showing MAPE tolerance for dataset of various stocks and the outcome of the current jump analysis method showing MAPE tolerance for dataset of various stocks studied by Hassan. Indices of close price from 1997 to 2004 as shown in Table 8 are used for verification of the daily TAIEX dataset. Table 8 summarizes various methods developed for forecasting the TAIEX index. The performance is compared with that of the existing models as shown in Fig. 5. The data from January to October for each year have been used as training dataset and November, December data for forecasting. The first 10 months have been used as training the model and the last two months for testing the model. Extensive analysis has been carried out for various stock indices with a multi-time-frame approach to ensure it as the best effective model predicting the trend. There is considerable uncertainty in decision making problems. But the current model helps in successful prediction of patterns, trends using error tolerance limits giving optimum solutions.

Conclusion

A stock market moving trend is a tendency of a financial market to move in a specific direction over time. Buyers and sellers move markets based on expectations and emotions (fear and greed). In this current study, a NTFCMs interrelated model and integrated NTFCMs Markov chain model in which hidden

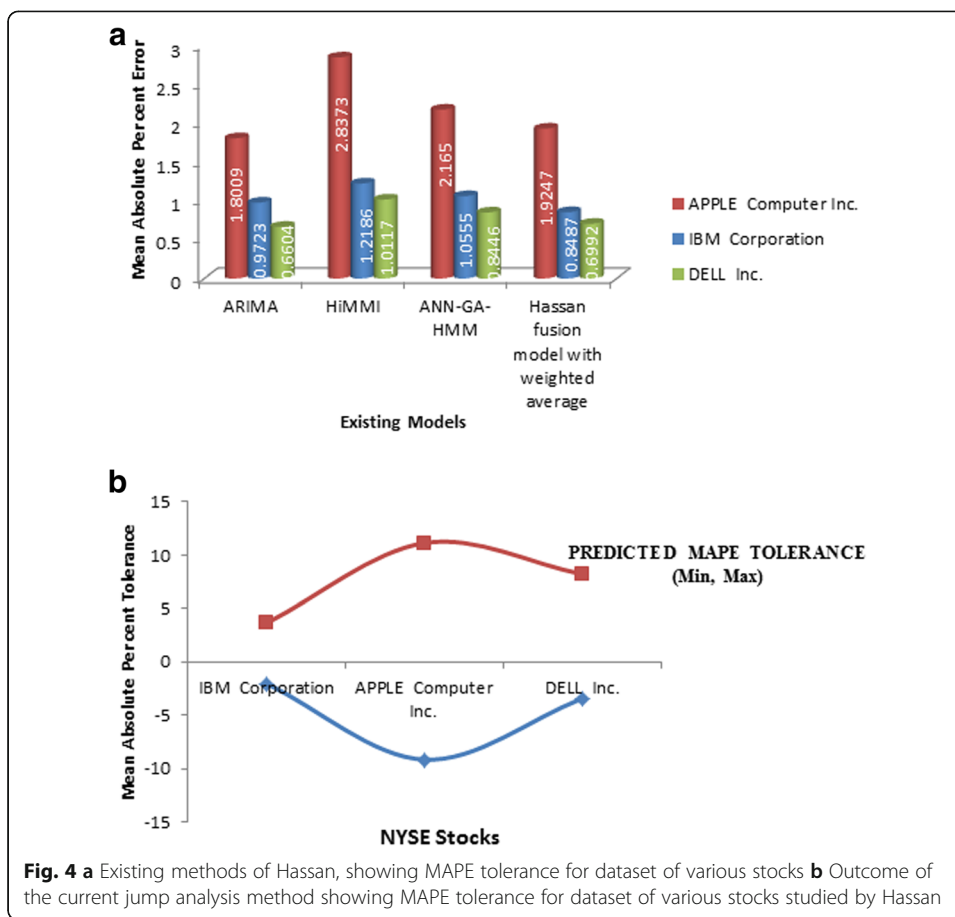
Table 7 Comparison of the outcome of the Markov chain and fuzzy interrelated method with other existing methods

Stock Name	Mean absolute per cent error (MAPE) in forecast for test dataset					Observed trend	Existing trend	Error
	ARIMA	HIMMI	ANN-GA-HMM Interpolation	Hassan fusion model with weighted average	The proposed model MAPE tolerance limit			
IBM Corp.	0.9723	1.2186	1.0555	0.8487	(−2.03, 3.68)	S/M trend	S/M trend	Zero
APPLE Comp. Inc.	1.8009	2.8373	2.1650	1.9247	(−9.15, 11.14)	Uptrend	Uptrend	Zero
DELL Inc.	0.6604	1.0117	0.8446	0.6992	(−3.42, 8.22)	S/M trend	S/M trend	Zero

Train dataset collected from 10.2.2003 to 10.9.2004

Test dataset collected from 13.9.2004 to 21.1.2005

S/M Sideways or mixed

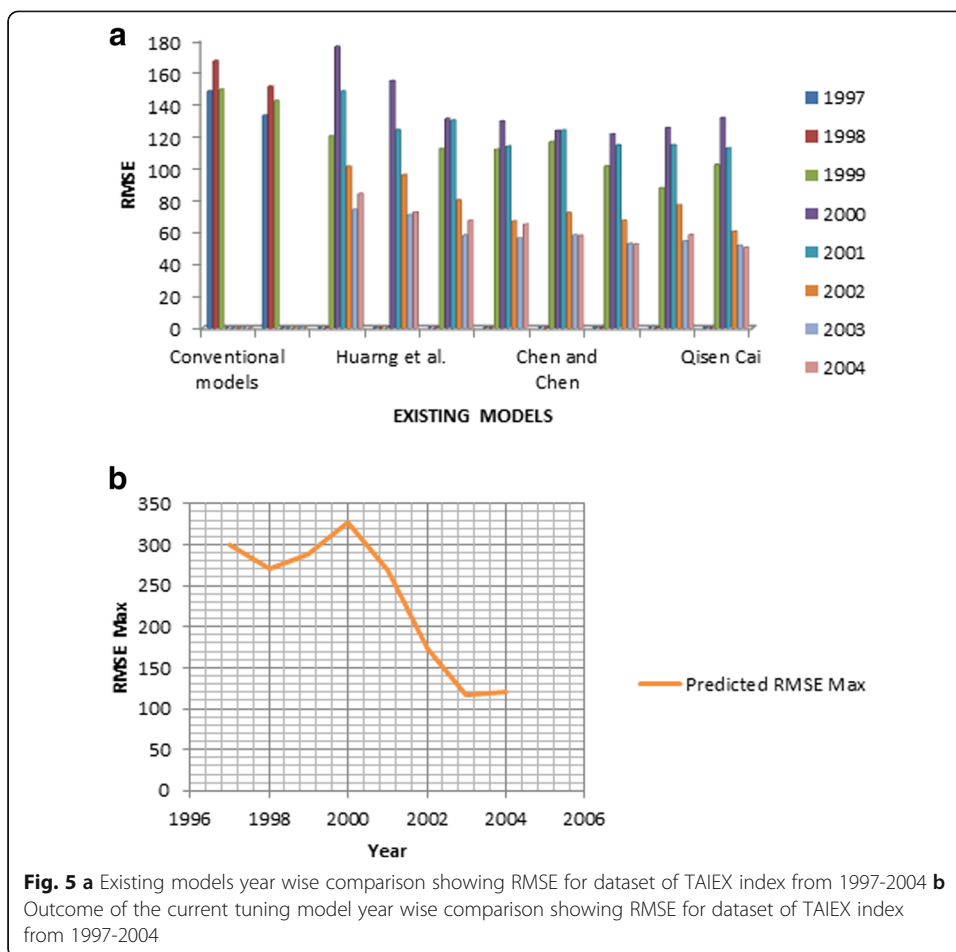


threshold values have been used in the very construction of the nodes of the dynamic system of the FCMs. The study of the movements in the stock markets is based upon the use of this model. The current model has helped with the prediction of the trend indicating its lower bound, upper bound and the exact threshold values giving the best expected most optimal solution. The main objective is to illustrate the efficacy of the current model with extensive analysis of three different stocks namely IBM, UNG, APPLE INC which followed an upward trend, a downward trend and a mixed trend respectively. The contribution of the current research work is finding the error tolerance limit for various errors like MAPE, RMSE errors for various stock indices using a multi-timeframe approach. The predicted minimum and predicted maximum of the errors help in enhancement of the decision making process for achievement of reliable forecasts. All the existing methods reveal the error as being within the range of tolerance limits. Results showed the predicted value within tolerance limit for every 'n' number of predictions made with nil error and 100% accuracy. Even if many more tools and techniques are developed in future to enhance forecast accuracy, all those developed would still have an error within this tolerance range. There is no possibility for deviation from the defined range. Predicting the close price for next one day, one week, one month, two months, and four months ahead can never be so accurate

Table 8 Year wise comparison for TAIEX index from 1997 to 2004

	1997	1998	1999	2000	2001	2002	2003	2004
Conventional models	148	167	149	-	-	-	-	-
Weighted models	133	151	142	-	-	-	-	-
Yu and Huang	-	-	120	176	148	101	74	84
Huang et al.	-	-	N/A	154.42	124.02	95.73	70.76	72.35
Yu et al.	-	-	112	131	130	80	58	67
Chen and Chang	-	-	111.70	129.42	113.67	66.82	56.10	64.76
Chen and Chen	-	-	116.64	123.62	123.85	71.98	58.06	57.73
Chen et al.	-	-	101.33	121.27	114.48	67.18	52.72	52.27
Chen and Kao	-	-	87.63	125.34	114.57	76.86	54.29	58.17
Qisen Cai	-	-	102.22	131.53	112.59	60.33	51.54	50.33
The proposed model	(0, 300.72)	(0, 269.93)	(0, 287.83)	(0, 326.31)	(0, 268.35)	(0, 173.46)	(0, 115.98)	(0, 119.68)
tolerance limit	Zero	Zero	Zero	Zero	Zero	Zero	Zero	Zero
Error	(-5.14, 3.35)	(-2.90, 3.53)	(-4.15, 3.43)	(-9.21, 5.54)	(-5.00, 5.19)	(-3.11, 3.46)	(-2.21, 1.75)	(-4.01, 2.16)
Predicted MAPE	S/M trend	S/M trend	S/M trend	Down trend	Up trend	S/M trend	Down trend	S/M trend
Predicted Trend								

S/M Sideways or mixed



and satisfactory. We therefore ensure accurate prediction of the range of tolerance predicting the minimum and maximum error giving an optimum solution and thereby enhancing successful trend prediction model. The future research work and plan is to develop a system for improvement of accuracy in the error tolerance limit.

Abbreviations

NTFCMs: New Threshold Fuzzy Cognitive Maps; MAPE: Mean Absolute Percentage Error; RMSE: Root mean square error; TAIEX: Taiwan Futures Exchange; HMM: Hidden Markov Model; ANN: Artificial Neural Network; GA: Genetic Algorithm; FCMs: Fuzzy cognitive maps; IBM: International Business Machines Corporation

Acknowledgements

‘Not applicable’

Authors’ contributions

GK helped to conceive and design the study, conducted the data collection and analysis and drafted the manuscript. AU helped to conceive and design the study, coordinated the research activities. DN helped to draft and review the manuscript. All authors read and approved the final manuscript.

Funding

‘Not applicable’.

Availability of data and materials

The stock prices for IBM, APPLE INC, UNG sectors at the close of every day for a period of one month are collected from www.finance.yahoo.com. The authors declare that they have no competing interests.

Competing interests

The authors declare that they have no competing interests.

Received: 13 December 2018 Accepted: 8 August 2019

Published online: 29 August 2019

References

- Cai Q, Zhang D, Zheng W, Stephen CH, Leung (2015) A new fuzzy time series forecasting model combined with ant colony optimization and auto-regression. *Knowl-Based Syst* 74:61–68
- Ceska M, Jansen N, Junges S, Kataoen J-P (2019) Shepherding hordes of Markov chains. In: *Advances in knowledge discovery and data mining*. <https://doi.org/10.1007/978-3-030-17465-1-10>
- Chao X, Kou G, Peng Y, Alsaadi FE (2019) Behavior monitoring methods for trade-based money laundering integrating macro and micro prudential regulation: a case from China. *Technol Econ Dev Econ*:1–16. <https://doi.org/10.3846/tede.2019.9383>
- Chen MY, Chen BT (2015) A hybrid fuzzy time series model based on granular computing for stock price forecasting. *Inform Sci* 294:227–241
- Chen SM, Chang YC (2010) Multi-variable fuzzy forecasting based on fuzzy clustering and fuzzy rule interpolation techniques. *Inform Sci* 180:4772–4783
- Chen SM, Chen CD (2011) TAIEX forecasting based on fuzzy time series and fuzzy variation groups. *IEEE Trans Fuzzy Syst* 19:1–12
- Chen SM, Chu HP, Sheu TW (2012) TAIEX forecasting using fuzzy time series and automatically generated weights of multiple factors. *IEEE Trans Syst Man Cybern Part A Syst Hum* 42:1485–1495
- Chen SM, Kao PY (2013) TAIEX forecasting based on fuzzy time series. Particle swarm optimization techniques and support vector machines. *Inform Sci* 247:62–71
- Cheng YC, Li ST (2012) Fuzzy time series forecasting with a probabilistic smoothing hidden Markov model. *IEEE Trans Fuzzy Syst* 20:291–304
- Crone SF, Hibon M, Nikolopoulos K (2011) Corrigendum to ad-vances in forecasting with neural networks? Empirical evidence from the NN3 competition on time series prediction. *Int J Forecast* 27:635–660
- Denkena B, Nemeti A (2013) Stock market related pricing mechanisms for the tool and mould manufacturing industry. In: *Procedia CIRP* 12, eighth CIRP conference on intelligent computing in manufacturing engineering, pp 414–419
- Gharehchopogh FS, Bonab TH, Khaze SR (2013) A linear regression approach to prediction of stock market trading volume: a case study. *Int J Manag Value Supply Chains* 4(3):25
- Gray SA, Gray S, de Kok J-L et al (2015) Using fuzzy cognitive mapping as a participatory approach to analyze change, preferred states, and perceived resilience of social-ecological systems. *Ecol Soc* 20(2). <https://doi.org/10.5751/ES-07396-200211>
- Guo Z, Wang H, Yang J, Miller DJ (2015) A stock market forecasting model combining two-directional two-dimensional principal component analysis and radial basis function neural network. *PLoS One* 10(4):e0122385. <https://doi.org/10.1371/journal.pone.0122385>
- Hassan MR, Nath B (2005) Stock market forecasting using hidden Markov model: a new approach. In: *Proc. of the 2005 5th international conference on intelligent systems, design and applications*, 0-7695-2286-06/05, IEEE
- Hassan MR, Nath B, Kirley M (2007) A fusion model of HMM, ANN and GA for stock market forecasting. *Expert Syst Appl* 33(1): 171–180
- Hassan R (2009) A combination of hidden Markov model and fuzzy model for stock market forecasting. *Neurocomputing* 72: 3439–3446
- Hassan S, Jaafar J, Samir BB, Jilani TA (2012) A hybrid fuzzy time series model for forecasting. *Eng Lett* 20:1 EL-20-1-11
- Huang KH, Yu THK, Hsu YW (2007) A multivariate heuristic model for fuzzy time-series forecasting. *IEEE Trans Syst Man Cybern B Cybern* 37:836–846
- Hung J-C (2011) Applying a combined fuzzy systems and GARCH model to adaptively forecast stock market volatility. *Appl Soft Comput* 11:3938–3945
- Iyer S, Kamdar NR, Soparkar B (2015) Stock market prediction using digital signal processing models. *Int J Comput Appl* 6:35–39
- Kosko B (1986) Fuzzy Cognitive Maps. *Int J Man Mach Stud* 24:65–75
- Kou G, Chao X, Peng Y, Alsaadi FE, Herrera-Viedma E (2019) Machine learning methods for systemic risk analysis in financial sectors. *Technol Econ Dev Econ*:1–27. <https://doi.org/10.3846/tede.2019.8740>
- Li G, Kou G, Peng Y (2016) A group decision making model for integrating heterogeneous information. *IEEE Trans Syst Man Cybern Syst Hum* PP(99):1–11. <https://doi.org/10.1109/TSMC.2016.2627050>
- Li S-T, Kuo S-C, Cheng Y-C, Chen C-C (2010) Deterministic vector long-term forecasting for fuzzy time series. *Fuzzy Sets Syst* 161:1852–1870
- Madhoushi M, Aliabadi AN (2012) Forecasting stock exchange market using hybrid neuro fuzzy model. In: *The sixth international conference on fuzzy Informations and engineering*, pp 1–7
- Makridakis S, Hibon M (2000) The M3-competition: c xv results, conclusions and implications. *Int J Forecast* 16:451–476
- Park S, Lee J, Son Y (2016) Predicting market impact costs using nonparametric machine learning models. *PLoS One* 11(2): e0150243. <https://doi.org/10.1371/journal.pone.0150243>
- Pritpal S, Bhogeswar B (2013) High-order fuzzy-neuro expert system for time series forecasting. *Knowl-Based Syst* 46:12–21
- Rabiner L (1993) A tutorial on HMM and selected applications in speech recognition. *Proc IEEE* 77:267–296
- Rabiner L, Juang B (1993) *Fundamentals of speech recognition*. Prentice-Hall, Englewood Cliffs
- Rafiq H, Baikunth N, Kirley M (2006) HMM based fuzzy model for time series prediction. In: *IEEE international conference on fuzzy systems*, pp 2120–2126
- Vasantha Kandasamy WB, Smarandache F (2003) *Book on fuzzy cognitive maps and Neutrosophic cognitive maps*. American Research Press, Xiquan. arXiv:math/0311063[math.GM].
- Wahab AF, Zulkij M'd IE, Rahim HA, Akaria R (2013) Interval type-2 fuzzy logic system model in measuring the index value of underground economy in Malaysia. *Appl Math Sci* 102:5011–5084
- Wei LY, Cheng CH, Wu HH (2014) A hybrid ANFIS based on n-period moving average model to forecast TAIEX stock. *Appl Soft Comput* 19:86–92

- Weron R (2014) Electricity price forecasting: a review of the state-of-the-art with a look into the future. *Int J Forecast* 30:1030–1081
- Xue Y, Liu Z, Luo J, Ma Z, Zhang M, Hu X, Kuang Q (2015) Stock market trading rules discovery based on biclustering method. *Math Probl Eng* 2015:13. Article ID 849286. <https://doi.org/10.1155/2015/849286>.
- Yu HK (2005) Weighted fuzzy time series models for TAIEX forecasting. *Physica A* 349:609–624
- Yu THK, Huarng KH (2008) A bivariate fuzzy time series model to forecast the TAIEX. *Expert Syst Appl* 34:2945–2952
- Yu THK, Huarng KH (2010) Corrigendum to “A bivariate fuzzy time series model to forecast the TAIEX”. [*Expert Systems with Applications* 34 (4), 2945–2952 (2010)]. *Expert Syst Appl* 37:5529
- Zhang H, Kou G, Peng Y (2019) Soft consensus cost models for group decision making and economic interpretations. *Eur J Oper Res* 277(3):964–980
- Zhang W, Zhang X, Sun Y (2017) A new fuzzy cognitive map learning algorithm for speech emotion recognition. *Math Probl Eng* 2017. <https://doi.org/10.1155/2017/4127401>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
