Applied Soft Computing xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Applied Soft Computing

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

A hybrid ANFIS model based on empirical mode decomposition for stock time series forecasting

Q1 Liang-Ying Wei *

Department of Information Management, Yuanpei University of Medical Technology, 306 Yuanpei Street, Hsin Chu 30015, Taiwan

ARTICLE INFO

Article history:
Received 17 August 2012
Received in revised form 19 January 2016
Accepted 19 January 2016
Available online xxx

Keywords:
Adaptive network-based fuzzy inference system (ANFIS)
Empirical mode decomposition (EMD)
Stock forecasting

ABSTRACT

Time series forecasting is an important and widely popular topic in the research of system modeling, and stock index forecasting is an important issue in time series forecasting. Accurate stock price forecasting is a challenging task in predicting financial time series. Time series methods have been applied successfully to forecasting models in many domains, including the stock market. Unfortunately, there are 3 major drawbacks of using time series methods for the stock market: (1) some models can not be applied to datasets that do not follow statistical assumptions; (2) most time series models that use stock data with a significant amount of noise involutedly (caused by changes in market conditions and environments) have worse forecasting performance; and (3) the rules that are mined from artificial neural networks (ANNs) are not easily understandable.

To address these problems and improve the forecasting performance of time series models, this paper proposes a hybrid time series adaptive network-based fuzzy inference system (ANFIS) model that is centered around empirical mode decomposition (EMD) to forecast stock prices in the Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) and Hang Seng Stock Index (HSI). To measure its forecasting performance, the proposed model is compared with Chen's model, Yu's model, the autoregressive (AR) model, the ANFIS model, and the support vector regression (SVR) model. The results show that our model is superior to the other models, based on root mean squared error (RMSE) values.

1. Introduction

Q4 System modeling is an intelligent modeling approach for establishing a target system for function approximation, optimization, and forecasting and has been examined extensively for many years. Time series forecasting is an important application in system modeling. In a time series, time is usually an important variable in making decisions or forecasts. Managers usually use historical data to forecast various types of variables, such as changes in stock price and product sales. The use of time series to forecast future tendencies has to make use of detailed data that were generated some time in the past to understand changes in trends.

Forecasting stock prices is an important topic in finance and a good application for demonstrating the value of time series forecasting. Successful forecasting of stock prices presumably results in substantial monetary rewards. Modern business and economic activities, in essence, are dynamic and change frequently and financial time series show nonlinearity and nonstationary behavior. Thus, it is difficult to accurately forecast the movements of stock prices, because stock markets are affected by many intricately related economic and political factors.

Many traditional time series models have been proposed and applied to economic forecasting. Engle [7] developed the ARCH (p) (autoregressive conditional heteroscedasticity) model, which has been used by many financial analysts; the GARCH [1] (generalized ARCH) model is the generalized form of ARCH. Box and Jenkins [2] established the autoregressive moving average (ARMA) model, which combines a moving average with a linear difference equation and makes forecasts under linear stationary conditions.

Models that describe such homogeneous nonstationary behavior can be obtained by assuming suitable differences in the process that is to be stationary. To this end, the autoregressive integrated moving average model (ARIMA) [2], with the assumption of linearity between variables, was proposed to handle nonstationary behavior datasets. However, the conventional time series method requires more historical data, and the data must have a normal distribution to effect better forecasting. Further, linguistic expressions are often used to describe daily observations.

Thus, Song and Chissom [29] proposed the original fuzzy time series model, and subsequent researchers focused on the

Q2 * Tel.: +886 936244363.
E-mail address: lywei@yuntech.edu.tw

http://dx.doi.org/10.1016/j.asoc.2016.01.027
1568-4946/

2 major processes of the fuzzy time series model: (1) fuzzification and (2) establishment of fuzzy relationships and forecasting. In the fuzzification process, the length of intervals for the universe of discourse can affect the forecast, and Huarng proposed a distribution-based and average-based length to address this issue [11]. In addition, Chen proposed a method [5] in which the length of linguistic intervals is tuned by genetic algorithms. In establishing fuzzy relationships and forecasting, Yu [35] argued that recurrent fuzzy relationships should be considered in forecasting and recommended that different weights be assigned to various fuzzy relationships. Thus, Yu [35] introduced a weighted fuzzy time series method to forecast the TAIEX. To take advantage of neural networks (nonlinear capabilities), Huarng and Yu [15] chose a neural network to establish fuzzy relationships in fuzzy time series, which are also nonlinear, but the process of mining fuzzy logical relationships is not easily understood [3].

Artificial intelligence (AI) is a rapidly growing technology with regard to information processing applications. It has been applied to various disciplines, such as business, engineering, management, science, and the military. In the financial domain, AI can be used to predict stock prices, credit scores, and potential bankruptcy problems. Stock prices change dramatically, which impedes the accurate prediction of stock price volatility.

In the past decade, many studies have been conducted to mine financial time series data, including traditional statistical approaches and data-mining techniques. Many researchers have applied data-mining techniques to financial analysis [12,19–23,25,27]. Further, many hybrid forecasting techniques have been published recently [9,10,24,26,28]. ANFIS [17] (an AI method) is a hybrid technique that integrates the advantage of learning in an ANN and using a set of fuzzy if-then rules with appropriate membership functions to generate input-output pairs with a high degree of accuracy [17]. In recent years, that ANFIS system has been used widely to generate nonlinear models of processes to determine input-output relationships. Thus, ANFIS is appropriate for forecasting nonlinear financial time series and generating meaningful rules for strategizing investment tactics.

The research topic of this paper is financial time series, the characteristics of which, such as its nonlinear or nonstationary nature and the presence of noises in the raw data, should be considered. However, traditional time series models have linear limitations and are unsuitable for financial forecasting. Thus, hybrid models are widely used to circumvent the limitations in financial time series forecasting. Empirical mode decomposition (EMD) [14] is perfectly suitable for nonlinear and nonstationary time series signal analysis [14] and identifies tendencies in financial time series. Based on EMD, any complicated signal can be decomposed into a finite and often small number of intrinsic mode functions (IMFs) [14], which have simpler frequency components and stronger correlations, rendering them easier to forecast more accurately. EMD has been used widely in many fields, such as in the analysis of earthquake signals, structure analysis, bridge and construction monitoring [34], sea wave data [13], and the diagnosis of faults in machines [36].

Based on the findings above, there are 3 major drawbacks of these models: (1) For certain statistical models, specific assumptions are required for observations, and the models cannot be applied to datasets that do not follow them. (2) Most conventional time series models use late-day stock price as the input variable in forecasting. However, there is a significant amount of involuted noise in raw stock data that is caused by changes in market conditions and environments. Traditional time series models that use complicated raw data certainly have reduced forecasting performance. For this reason, forecasting models should decompose complicated raw data into simpler frequency components and high-correlation variables to improve their forecasting accuracy. (3) ANN is a black box method, and the rules mined from ANNs are not easily understandable [3]. Nevertheless, forecasting rules are useful for investors in buying and selling stocks.

To overcome the drawbacks above, this paper hypothesizes that EMD can decompose complicated raw data (stock index) into simpler frequency components and highly correlating variables, adopting it into an AR model to build the primary model. Then, the results of the primary model are refined and optimized by an adaptive network-based fuzzy inference system (ANFIS) that uses fuzzy if-then rules to model the qualitative aspects of human knowledge for applicability to human activities.

Based on the concepts that have been discussed, this paper proposes a hybrid ANFIS model that considers stock index (t) to forecast future stock index values ($t+1$). First, this approach uses EMD to decompose the original stock index (t) data into IMFs (intrinsic mode functions) and a residue (R). Then, the tendencies of these IMFs and the residue are modeled and forecasted using ANFIS, which can overcome the limitations of statistical methods (the data need to obey some mathematical distribution). Finally, the prediction results are integrated to obtain a final forecasting value. Thus, this study expects that the proposed model can generate significant profits for investors by more accurately forecasting the stock market.

The remainder of the paper is organized as follows. Section 2 describes the related studies. Section 3 briefly presents the proposed model. Section 4 discusses the experiments and makes comparisons. Section 5 presents the findings of the experiment results. Finally, conclusions are made in Section 6.

2. Related works

This section reviews the relevant studies on various forecasting models for the stock market, empirical mode decomposition, subtractive clustering (Subclust), and ANFIS.

2.1. Various forecasting models for the stock market

Participating in the stock market is one of the most exciting and challenging monetary activities. Market climates change dramatically in seconds. Thus, many methods have been developed to address problems with forecasting stock prices. Huarng et al. [16] used the volatility of the NASDAQ (the largest US electronic stock market) index and Dow Jones Industrial Average index to forecast the Taiwan stock index.

Time series models have been applied to make economic forecasts, such as stock index forecasting, for which various models have been proposed. Engle [7] introduced the ARCH (autoregressive conditional heteroscedasticity) model, which has been used by many financial analysts, and the GARCH [1]. Box and Jenkins [2] proposed the autoregressive moving average (ARMA) model, which combines a moving average with a linear difference equation. In recent years, many researchers have applied data-mining techniques to financial analysis. Huarng and Yu [15] applied the backpropagation neural network to establish fuzzy relationships in fuzzy time series to forecast stock price. Kinoto et al. [20] developed a prediction system for the stock market by using a neural network. Nikolopoulos and Fellrath [25] combined genetic algorithms and a neural network to form a hybrid expert system for investment advising. Kim and Han [19] introduced a genetic algorithm-based approach for feature discretization and the determination of connection weights for ANNs to predict stock price indices. Roh [27] integrated a neural network and a time series model to forecast the volatility of stock price indices. Thawornwong and Enke [31] proposed redeveloped neural network models to predict the directions of future excess stock returns. Kim [18] applied SVM (support vector machine) to predict stock price indices.

Please cite this article in press as: L.-Y. Wei, A hybrid ANFIS model based on empirical mode decomposition for stock time series forecasting, Appl. Soft Comput. J. (2016), http://dx.doi.org/10.1016/j.asoc.2016.01.027

2.2. Empirical mode decomposition

The empirical mode decomposition (EMD) technique, proposed by Huang et al. [14], is an adaptive time series decomposition technique that uses the Hilbert-Huang transform (HHT) for nonlinear and nonstationary time series data. The basic principle of EMD is to decompose a time series into a sum of oscillatory functions—namely, intrinsic mode functions (IMFs). In EMD, IMFs must satisfy 2 conditions: (1) the number of extrema (sum of maxima and minima) and the number of 0 crossings must differ only by 1, and (2) the local average is 0. The latter condition implies that the envelope mean of the upper and lower envelopes is equal to 0. The first condition is similar to the traditional narrow band requirements for a stationary Gaussian process [14]. The second condition modifies the classical global requirement to a local one, so that the instantaneous frequency will not experience unwanted fluctuations due to asymmetric wave forms [14]. IMFs are naturally ordered according to their decreasing frequency content and end with a nonoscillatory trend. The process of creating intrinsic modes by an iterative procedure is called sifting.

During the sifting mechanism, an intermediate signal $r(t)$, created initially with a copy of the original time series $r(t) = x(t)$, is determined as follows:

1. Find local extremes of $r(t)$.
2. Compute 2 envelopes using the local maxima and local minima, respectively. Various interpolation functions can be used. In this study, a cubic spline interpolation scheme is used.
3. The average $m(t)$ of the 2 envelopes is calculated and subtracted from $r(t)$, yielding $h(t) = r(t) - m(t)$.
4. If $h(t)$ is not an IMF, then repeat the sifting after updating $r(t) = h(t)$. Otherwise, set $c_j(t) = h(t)$, and estimate $j+1$ – th IMF after updating $r(t) = x(t) - \sum_{j<(j+1)} c_j(t)$.

This procedure stops when a nonoscillatory residual signal is encountered and all IMFs are extracted. Alternatively, the number of modes that are to be extracted can be fixed deliberately by the user. The entire procedure of decomposing a time series via EMD is explained schematically in Fig. 1. EMD also ensures the perfect reconstruction property—i.e., summing up all of the extracted IMFs with the residual trend reconstructs the original signal without loss or distortion of information. This also implies that the amplitudes of the extracted components reflect the contribution of the component to the original signal. There is no scale indeterminacy, as with an ICA decomposition, for example.

Ultimately, the original signal $x(t)$ can be expanded into its intrinsic modes $c_j(t)$ and a non-oscillating residual signal $r(t)$ per Eq. (1):

$$x(t) = \sum_{j=1}^{j} c_j(t) + r(t) = \sum_{j=1}^{j+1} c_j(t) \tag{1}$$

where $c_j(t)$ represent the IMFs and $r(t)$ is the remaining nonoscillating trend. For the last equality, $c_{j+1}(t) = r(t)$ has been taken.

2.3. Subtractive clustering

Chiu [6] developed subtractive clustering, a fuzzy clustering method, to estimate the number and initial locations of cluster centers. Consider a set T of N data points in a D-dimensional hyperspace, with each data point W_i ($i = 1, 2, ..., N$). $W_i = (x_i, y_i)$, where x_i denotes the p input variables and y_i is the output variable. The potential value P_i of the data point is calculated by Eq. (2):

$$P_i = \sum_{j=1}^{N} e^{-\alpha ||W_i - W_j||^2} \tag{2}$$

where $\alpha = 4/r^2$, r is the radius that defines a W_i neighborhood, and $||.||$ denotes the Euclidean distance.

The data point with many neighboring data points is chosen as the first cluster center. To generate the other cluster centers, the potential P_i is revised for each data point W_i per Eq. (3)

$$p_i = p_i - p_1^* \exp(-\beta ||W_i - W_1^*||^2) \tag{3}$$

where β is a positive constant that defines the neighborhood that will have measurable reductions in potential. W_1^* is the first cluster center, and P_1^* is its potential value.

From Eq. (3), the method selects the data point with the highest remaining potential as the second cluster center. For the general equation, we can rewrite Eq. (3) as Eq. (4):

$$p_i = p_i - p_k^* \exp(-\beta ||W_i - W_k^*||^2) \tag{4}$$

where $W_k^* = (x_k^*, y_k^*)$ is the location of the k'th cluster center and P_k^* is its potential value.

At the end of the clustering process, the method obtains q cluster centers and D corresponding spreads S_i, $i = (1, \ldots, D)$. Then, we define their membership functions. The spread is calculated according to β.

2.4. ANFIS: adaptive network-based fuzzy inference system

Jang [17] proposed the adaptive network-based fuzzy inference system (ANFIS), a fuzzy inference system that is implemented in the framework of adaptive networks. To illustrate the system, we assume a fuzzy inference system that comprises 5 layers of an adaptive network with 2 inputs x and y and 1 output z. The architecture of ANFIS is shown in Fig. 2.

Then, we suppose that the system consists of 2 fuzzy if-then rules, based on Takagi and Sugeno's type [30]:

Rule 1: If x is A_1 and y is B_1, then $f_1 = p_1 x + q_1 y + r_1$
Rule 2: If x is A_2 and y is B_2, then $f_2 = p_2 x + q_2 y + r_2$.

The node in the i-th position of the k-th layer is denoted as $O_{k,i}$, and the node functions in the same layer are of the same function family, as described below:

Layer 1: This layer is the input layer, and every node i in this layer is a square node with a node function [see Eq. (5)]. $O_{1,i}$ is the membership function of A_i and specifies the degree to which the given x satisfies the quantifier A_i. Usually, we select a bell-shaped membership function as the input membership function [see Eq. (6)] with the maximum equal to 1 and the minimum equal to 0.

$$O_{1,i} = \mu_{A_i}(x) \quad \text{for} \quad i = 1, 2 \tag{5}$$

$$\mu_{A_i}(x) = \frac{1}{1 + [(x - c_i/a_i)^2]^{bi}} \tag{6}$$

where a_i, b_i, and c_i are the parameters, b is a positive value, and c denotes the center of the curve.

Layer 2: Every node in this layer is a square node, labeled Π, which multiplies the incoming signals and sends the product out per Eq. (7).

$$O_{2,i} = w_i = \mu_{A_i}(x) \times \mu_{B_i}(y) \quad \text{for} \quad i = 1, 2 \tag{7}$$

Layer 3: Every node in this layer is a square node, labeled N. The i-th node calculates the ratio of the i-th rule's firing strength to the

Please cite this article in press as: L.-Y. Wei, A hybrid ANFIS model based on empirical mode decomposition for stock time series forecasting, Appl. Soft Comput. J. (2016), http://dx.doi.org/10.1016/j.asoc.2016.01.027

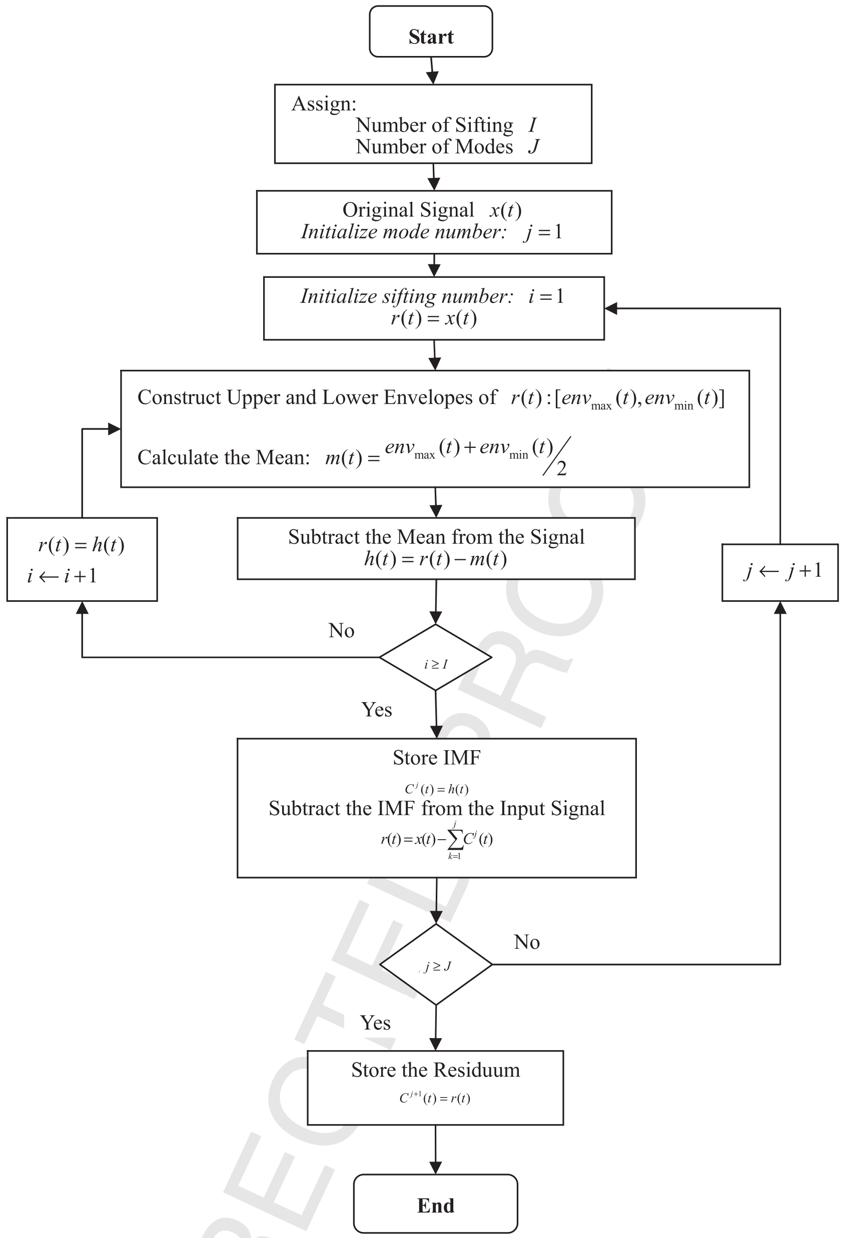

Fig. 1. Flowchart of the EMD algorithm.

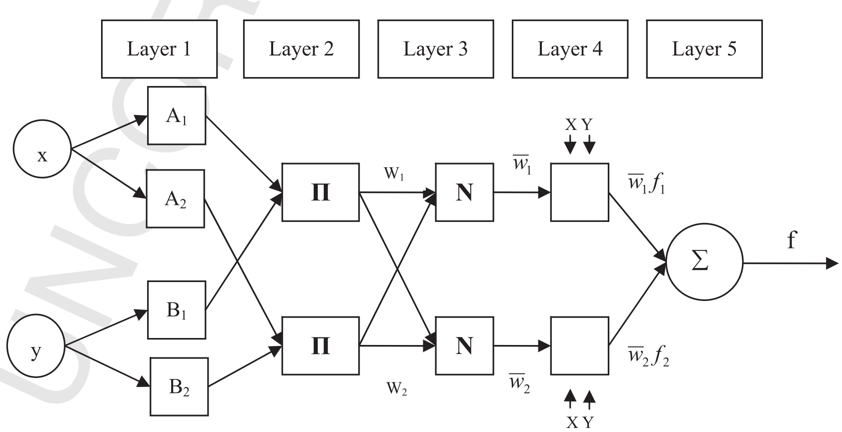

Fig. 2. The architecture of the ANFIS network.

sum of the firing strengths of all rules per Eq. (8). The output of this layer can be called normalized firing strengths.

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad \text{for} \quad i = 1, 2 \tag{8}$$

Layer 4: Every node i in this layer is a square node with a node function [see Eq. (9)]. Parameters in this layer will be referred to as consequent parameters:

$$O_{4,i} = \bar{w}_i f_i = \bar{w}_i (p_i + q_i + r_i) \tag{9}$$

where p_i, q_i, and r_i are the parameters.

Layer 5: The single node in this layer is a circle node, labeled Σ, that computes the overall output as the summation of all incoming signals [see Eq. (10)]:

$$O_{5,i} = \sum_i \bar{w}_i f_i = \frac{\sum_{i=1} w_i f}{\sum_{i=1} w_i} = \text{overall output} \tag{10}$$

3. Proposed model

Based on our literature review, there are three major drawbacks of stock forecasting models. To overcome these drawbacks, this paper hypothesizes that EMD can decompose noisy raw data (stock index) into simpler frequency components and highly correlating variables and adapts it to the AR model to build the primary model. Then, the results of primary model will be refined by ANFIS, which can overcome the limitations of statistical methods (data need obey some mathematical distribution) and handle noise data involutedly.

Based on the concepts above, this paper proposes a hybrid time series model that considers a stock index (t) to forecast stock index ($t+1$) prices, where t denotes the t-th day and stock-index (t) denotes the stock index at t days. First, this approach uses EMD to decompose the original stock index (t) into IMFs and a residue. Then, the proposed model uses the fuzzy inference system to forecast a stock index, which combines the AR model, the tendencies of these IMFs, and the residue to forecast TAIEX ($t+1$). Third, this study optimizes the fuzzy inference system parameters using an adaptive network. The flowchart of the proposed model is shown as Fig. 3.

This section uses practically collected data as an example and shows the core concept of the proposed algorithm stepwise as follows.

Step 1: Collect dataset

In this step, TAIEX data from 2000 to 2006 (7 subdatasets) and Hang Seng index (HSI) datasets from 2000 to 2004 (5 subdatasets) are collected to illustrate the proposed model. For each year, every training dataset is selected from January to October, and the remaining dataset (from November and December) is used for testing. For ease of understanding, this paper only uses TAIEX datasets to demonstrate the proposed model from Steps 2 to 6.

Step 2: Decompose input variable by EMD

To obtain interpretable information on the input variable [TAIEX(t)], we use EMD to decompose the input variable [TAIEX(t)] into a finite set of IMFs [the residual $r_n(t)$ is also considered an IMF]. In this study, TAIEX(t)is the input signal to be decomposed into 5 IMFs and 1 residue (just 6 IMFs in each year's dataset), exhibiting stable and regular variation. Thus, the interruption and coupling between characteristics in the original data have been weakened to an extent, rendering the forecasting model easier to build. Finally, the IMFs are the output signal that is generated in Step 2.

Step 3: Generate ANFIS forecast model

The ANFIS forecast model will be introduced in Steps 3.1 to 3.3, using subtractive clustering [6] to partition the universe of discourse for input variables and generating the fuzzy inference system. The substeps of Step 3 are described as follows:

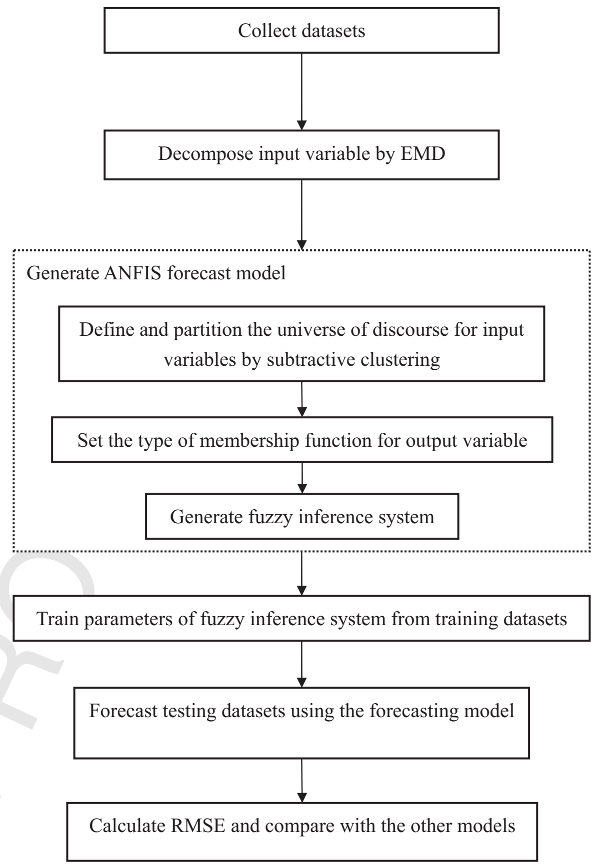

Fig. 3. Flowchart of the proposed procedure.

Step 3.1: Define and partition the universe of discourse for input variables by subtractive clustering

In this step, IMFs are the input signal, and the membership functions of ANFIS are generated by subtractive clustering.

First, the proposed model defines each universe of discourse for variables (IMFs) according to the minimum and maximum values for each variable (IMFs). The, subtractive clustering partitions the universe of discourse; the parameters for subtractive clustering are as follows: Gaussian membership function, range of influence = 0.8, accept ratio = 0.5, and reject ratio = 0.15. Each IMF is partitioned into 3 Gaussian membership functions (low, middle, and high). For example, the range of the Gaussian membership for $IMF_1(t)$, decomposed for the year 2000 TAIEX, ranges from −193.1 to 252.8.

The subtractive clustering method assumes that each data point is a potential cluster center and calculates a measure of the likelihood that each data point defines the cluster center, based on the density of surrounding data points. The algorithm:

1. Selects the data point with the highest potential to be the first cluster center.
2. Removes all data points in the vicinity of the first cluster center (as determined by *radii*) order to determine the next data cluster and its center location.
3. Iterates this process until all data are within *radii* of a cluster center.

The subtractive clustering method is an extension of the mountain clustering method that is proposed. The matrix X contains the data that are to be clustered; each row of X is a data point. The variable 'radii' is a vector of entries between 0 and 1 that specifies a cluster center's range of influence in each of the data dimensions, assuming that the data fall within a unit hyperbox. Small

Please cite this article in press as: L.-Y. Wei, A hybrid ANFIS model based on empirical mode decomposition for stock time series forecasting, Appl. Soft Comput. J. (2016), http://dx.doi.org/10.1016/j.asoc.2016.01.027

radii values generally result in a few large clusters. Good values for *radii* usually lie between 0.2 and 0.5.

Step 3.2: Set the type of membership function for the output variable

In this step, this paper sets a linear-type membership function for output variables [TAIEX $(t+1)$]. From Steps 2 and 3.1, there are 6 IMFs (including residual) that are decomposed from the input variable [$TAIEX(t)$] and 3 linguistic intervals that are partitioned by subtractive clustering in each IMF. Thus, a typical rule in a Sugeno fuzzy model is described as follows:

$$\text{If}\ \mathrm{u}\,(IMF_1(t)) = A_i,\ \mathrm{v}\,(IMF_2(t)) = B_i,\ \mathrm{w}\,(IMF_3(t)) = C_i,$$
$$\mathrm{x}\,(IMF_4(t)) = D_i,\ \mathrm{y}(IMF_5(t)) = E_i,\ \text{and}\ \mathrm{z}\,(IMF_6(t)) = F_i,$$
$$\text{then output is}\ f_i = m_i\,\mathrm{u} + n_i\,\mathrm{v} + o_i\mathrm{w} + p_i\,\mathrm{x} + q_i\,\mathrm{y} + r_i\,\mathrm{z} + s_i$$

where $\mathrm{u}(IMF_1(t))$, $\mathrm{v}(IMF_2(t))$, $\mathrm{w}(IMF_3(t))$, $\mathrm{x}(IMF_4(t))$, $\mathrm{y}\,(IMF_5(t))$, and $\mathrm{z}\,(IMF_6(t))$ are linguistic variables; A_i, B_i, C_i, D_i, E_i, and F_i are the linguistic labels (low, middle, high); f_i denotes the i-th output value; m_i, n_i, o_i, p_i, q_i, and r_i are the parameters; and s_i is constant (i = low, middle, high).

Step 3.3: Generate fuzzy inference system

First, the linguistic intervals, as input membership functions, are obtained in Step 3.1, and the output membership functions are set by Step 3.2. Then, the ANFIS model will generate fuzzy if-then rules, where the linguistic values ($A_i, B_i, C_i, D_i, E_i, F_i$) from the input membership functions are used as the 'if-condition' part, and the output membership function (f_i) is the 'then' component. The number of linguistic intervals that are partitioned by subtractive clustering is three (low, middle, and high), and the 3 generated rules (the number of rules is the same as the number of linguistic intervals) are described as follows:

Rule 1 (low):

$$\text{If}\ \mathrm{u}\,(IMF_1(t)) = A_{low},\ \mathrm{v}\,(IMF_2(t)) = B_{low},\ \mathrm{w}\,(IMF_3(t)) = C_{low},$$
$$\mathrm{x}\,(IMF_4(t)) = D_{low},\ \mathrm{y}\,(IMF_5(t)) = E_{low},\ \text{and}\ \mathrm{z}\,(IMF_6(t)) = F_{low},$$
$$\text{then output is}\ f_{low} = m_{low}\,\mathrm{u} + n_{low}\,\mathrm{v} + o_{low}\,\mathrm{w} + p_{low}\,\mathrm{x} + q_{low}$$
$$\mathrm{y} + r_{low}\mathrm{z} + s_{low}$$

Rule 2 (middle):

$$\text{If}\ \mathrm{u}\,(IMF_1(t)) = A_{middle},\ \mathrm{v}\,(IMF_2(t)) = B_{middle},\ \mathrm{w}\,(IMF_3(t)) = C_{middle},$$
$$\mathrm{x}\,(IMF_4(t)) = D_{middle},\ \mathrm{y}(IMF_5(t)) = E_{middle},\ \text{and}\ \mathrm{z}\,(IMF_6(t)) = F_{middle},$$
$$\text{then output is}\ f_{middle} = m_{middle}\,\mathrm{u} + n_{middle}\,\mathrm{v} + o_{middle}\mathrm{w} + p_{middle}\,\mathrm{x}$$
$$+ q_{middle}\,\mathrm{y} + r_{middle}\,\mathrm{z} + s_{middle}$$

Rule 3 (high):

$$\text{If}\ \mathrm{u}\,(IMF_1(t)) = A_{high},\ \mathrm{v}\,(IMF_2(t)) = B_{high},\ \mathrm{w}\,(IMF_3(t)) = C_{high},$$
$$\mathrm{x}\,(IMF_4(t)) = D_{high},\ \mathrm{y}\,(IMF_5(t)) = E_{high},\ \text{and}\ \mathrm{z}\,(IMF_6(t)) = F_{high},$$
$$\text{then output is}\ f_{high} = m_{high}\,\mathrm{u} + n_{high}\,\mathrm{v} + o_{high}\mathrm{w} + p_{high}\,\mathrm{x} + q_{high}\,\mathrm{y}$$
$$+ r_{high}\,\mathrm{z} + s_{high}$$

Step 4: Train fuzzy inference system parameters from training datasets

In this step, we use the least-squares method and back-propagation gradient descent method to train the forecasting model. This paper sets the stopping criterion to 1000 epochs, and the process is executed for the 1000 predetermined iterations unless it terminates while the training error converges. When the stopping criterion for the training set is reached, the optimal parameters for the selected output membership function are obtained. For example, the parameters for extracted rule 1 from the year 2000 TAIEX data are shown as follows:

$$m_{low}, n_{low}, o_{low}, p_{low}, q_{low}, r_{low}, s_{low} = -0.500, 0.677, 0.933, 1.014, 1.163, 1.025, -244$$

Step 5: Forecast testing datasets using the trained model

The fuzzy inference system parameters of the forecasting models are determined when the stopping criterion is reached from Step 4; then, the training forecasting model is used to forecast TAIEX$(t+1)$ for the target testing datasets.

Step 6: Calculate RMSE and compare with the other models

Calculate the RMSE values in the testing datasets per Eq. (11). Then, the RMSE is taken as the evaluation criterion to compare with the other models.

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{t=1}^{n} |actual(t) - forecast(t)|^2}{n}} \tag{11}$$

where $actual(t)$ denotes the real TAIEX value, $forecast(t)$ denotes the predicted TAIEX value, and n is the number of data points.

Table 1
The RMSE results of the models for TAIEX testing data.

Models	Year						
	2000	2001	2002	2003	2004	2005	2006
Chen	191	167	75	66	128	67	100
Yu	176	148	101	74	123	63	89
AR(1)	130	115	66	54	55.1	55 [a]	54
SVR	136	114	66	59	53	55 [a]	61
ANFIS	130	120	65	55	55	57	85
Proposed model	129[a]	110[a]	52[a]	49[a]	43[a]	82	42[a]

[a] The best performer among the 6 models.

4. Experiments and comparisons

In this section, we perform evaluations and make comparisons, using the RMSE as the evaluation criterion. To verify the proposed model, TAIEX datasets from 2000 to 2006 and Hang Seng index (HSI) datasets from 2000 to 2004 are used as the experiment datasets; each year of the TAIEX and HSI dataset is a subdataset.

Each subdataset for the previous 10 months is used for training, and those from November to December are selected for testing. Further, this paper compares the performance of the proposed model with that of the time series model, Chen's model [4], Yu's model [35], the AR (1) [7] model, and machine learning models, such as ANFIS [17] and SVR [33]. The performance of the other models above are compared with the proposed model, and Table 1 lists the comparison results for the TAIEX datasets. From Table 1, it is clear that the proposed model surpasses the other 5 models in each testing period, except for 2005.

Predicted values that were generated by the proposed model and the original TAIEX index for training datasets for the year 2000 are shown in Figs. 4 and 5 charts the predictions of the proposed model and the original values for the testing datasets in 2000.

Further, the comparison with the other models for the HSI datasets (from 2000 to 2004) is summarized in Table 2, which indicates that the proposed model outperforms the other models. Based on the results from Tables 1 and 2, these evaluations demonstrate the outstanding performance of the proposed model.

This paper uses a nonparametric statistical method, the Friedman test [8], to verify that the proposed model is superior to the other methods. Using the data from Tables 1 and 2, chi-square test is used to test the hypothesis H_0: equal performance. The results ($p = 0.001$ for TAIEX datasets, $p = 0.001$ for HSI datasets) with regard

Please cite this article in press as: L.-Y. Wei, A hybrid ANFIS model based on empirical mode decomposition for stock time series forecasting, Appl. Soft Comput. J. (2016), http://dx.doi.org/10.1016/j.asoc.2016.01.027

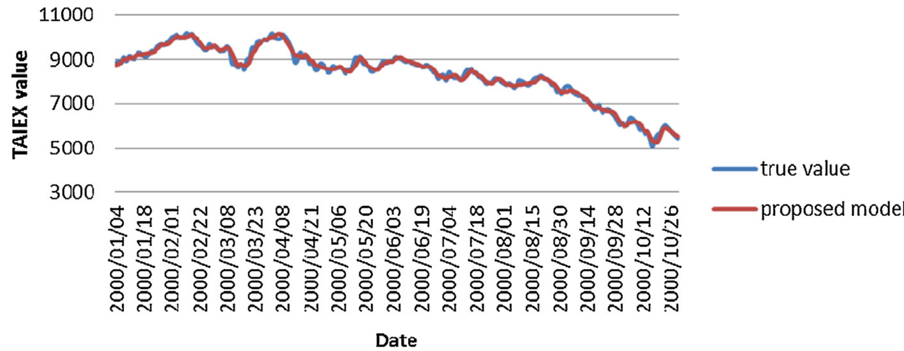

Fig. 4. The original and predicted values of TAIEX training datasets between January 4, 2000 and October 26, 2000.

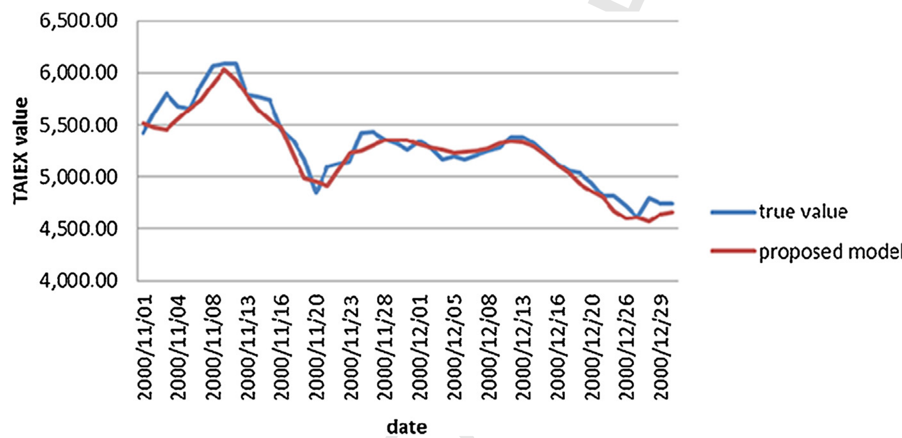

Fig. 5. The original and predicted values of TAIEX testing datasets between November 1, 2000 and December 29, 2000.

Table 2
The RMSE results of various models for HSI testing data.

Models	Year				
	2000	2001	2002	2003	2004
Chen	403	428	187	227	188
Yu	435	289	170	152	177
AR(1)	256	161	105	123	103
SVR	250	165	107	121	113
ANFIS	242	173	106	118	127
Proposed model	230[a]	147[a]	97[a]	78[a]	100[a]

[a] The best performer among the 6 models.

Table 3
Results of Friedman test (for TAIEX datasets).

Parameter	
n	7
Chi-Square	23.31
df	5
Asymp. Sig.	.001

Note: n is the number of data points; df denotes degree of freedom.

Table 4
Mean rank of Friedman Test (for TAIEX datasets).

Models	Mean Rank
Chen	5.57
Yu	5.14
AR(1)	2.64
SVR	2.86
ANFIS	3.07
Proposed model	1.71

Note: A smaller mean rank denotes high performance.

Table 5
Results of Friedman Test (for HSI datasets).

Parameter	
n	5
Chi-Square	21.8
df	5
Asymp. Sig.	.001

Note: n is the number of data points; df denotes degree of freedom.

to this hypothesis, which rejects $H_0 = 0$, are listed in Tables 3 and 5). Tables 4 and 6 show the mean rank by Friedman test for the TAIEX datasets and HSI, respectively, demonstrating that the proposed model (mean rank = 1.71 for TAIEX datasets, mean rank = 1 for HSI datasets) outperforms the other models for these datasets. Based on Tables 3–6, the difference in performance is significant.

For making advanced simulation trades and showing the resulting profits, this paper sets 2 trade rules for the Taiwan Futures Exchange (TAIFEX) to calculate profits and assumes that the profits unit is equal to 1. Thus, the profit formula is defined as Eq. (12).

Rule 1: sell rule

$$IF\ \frac{|forecast(t) - actual(t)|}{actual(t)} \le \alpha \quad \text{and} \quad forecast(t+1) - actual(t) > 0,$$

$$\text{then sell.}$$

Rule 2: buy rule

$$IF\ \frac{|forecast(t) - actual(t)|}{actual(t)} \le \alpha \quad \text{and,} \quad forecast(t+1)$$

$$-actual(t) < 0 \quad \text{then buy.}$$

Table 6
Mean rank of Friedman Test (for HSI datasets).

Models	Mean rank
Chen	5.8
Yu	5.2
AR(1)	2.8
SVR	3.2
ANFIS	3.0
Proposed model	1.0

Note: A smaller mean rank denotes high performance.

Table 7
Comparison of profits using the various models (TAIEX).

Year	α	Models					
		Chen's model	Yu's model	AR(1)	SVR	ANFIS	Proposed model
2000	0.01	−92	−73	671	202	686	795[a]

[a] The most profit among the 6 models.

where α denotes the threshold parameter ($0 < \alpha \leq 0.07$; the threshold parameter depends on daily fluctuations in the TAIEX).

Profit definition:

$$\text{Profit} = \sum_{t_s=1}^{p}(actual(t+1) - actual(t)) + \sum_{t_b=1}^{q}(actual(t) - actual(t+1)) \quad (12)$$

where p represents the total number of days for selling, q represents the total number of days for buying, t_s represents the t-th day for selling, and t_b represents the t-th day for buying.

The optimal threshold parameter α is obtained when the forecasting performance obtains the best profits in the training dataset. From the optimal threshold parameter α and Eq. (12), the profits for the various models are calculated, and the profits are shown in Table 7.

5. Findings

Based on our verification and comparison, the proposed method outperforms other methods, except for the 2005 TAIEX and HSI datasets, rendering it superior. By examining the performance of these models, there are 2 important findings as follows:

(1) **The advantage of the hybrid model**:
According to Tables 1 and 2, the proposed model is superior to the other methods in terms of RMSE, primarily because it takes into account the EMD method with ANFIS learning for stock index forecasting, integrating the advantage of ANFIS, which optimizes fuzzy inference system parameters using an adaptive network to improve forecasting performance.

(2) **EMD strength**:
From Tables 1 and 2, the proposed model performs better than the ANFIS model. It is evident that EMD can also reveal the hidden patterns and trends of time series, which can effectively assist financial time series forecasting.

6. Conclusions

This paper has developed a stock forecasting model by integrating EMD and ANFIS. The main contribution of the paper is that it proposes a novel method and a simple approach for making stable predictions of fluctuating data. The proposed method preprocesses stock index (t) and decomposes the index into more stationary and regular components (IMF or residue) using the EMD technique. Further, the corresponding ANFIS model for each divided component is easier to build. After the IMF components and residue are used to establish the ANFIS model, the forecasting values are the stock index ($t+1$) forecasting results. This study compared the proposed model with Chen's model; Yu's model; and the ANFIS, SVR, and AR (1) models, using RMSE as comparison criteria. The experimental results showed that the proposed EMD-ANFIS model produces the lowest forecasting errors and outperforms the other models for TAIEX and HSI datasets. In this study, EMD (a time-frequency resolution approach) offers a new approach by which the nonstationary and nonlinear behavior of time series can be decomposed into a series of valuable independent time resolutions and can reveal the hidden patterns of time series. Through this preprocessing, this study advances the simplification of ANFIS modeling and generates more precise forecasting results compared with traditional time series models that are based on RMSE. Thus, the proposed method is suitable for making predictions with nonlinear and high-noise data, such as time series datasets, and is an efficient method for stock index forecasting. Further, 2 advantages of the proposed model were discovered: (1) the proposed model can produce more reasonable and understandable rules, because the "if-then" rules that are produced by ANFIS can model the qualitative aspects of human knowledge; and (2) the proposed model can provide stock investors with objective suggestions (forecasts) to make investment decisions in the stock market, because it generates forecasting rules that are based on objective stock data rather than subjective human judgments.

The hybrid model should be tested in other stock markets, such as China and on individual stocks in Japan to examine its robustness. Other important input variables, such as technology indicators, could be combined in this model to enhance its accuracy, and different rule-based artificial intelligence techniques could be used to establish appropriate decision rules for investors.

Uncited reference Q5

[32].

References

[1] T. Bollerslev, Generalized autoregressive conditional heteroscedasticity, J. Econom. 31 (1986) 307–327.
[2] G. Box, G. Jenkins, Time Series Analysis: Forecasting and Control, Holden-Day, San Francisco, 1976.
[3] T.L. Chen, C.H. Cheng, H.J. Teoh, High-order fuzzy time-series based on multi-period adaptation model for forecasting stock markets, Phys. A 387 (2008) 876–888.
[4] S.M. Chen, Forecasting enrollments based on fuzzy time-series, Fuzzy Sets Syst. 81 (1996) 311–319.
[5] S.M. Chen, N.Y. Chung, Forecasting enrollments using high-order fuzzy time series and genetic algorithms, Int. Intell. Syst. 21 (2006) 485–501.
[6] S.L. Chiu, Fuzzy model identification based on cluster estimation, J. Int. Fuzzy Syst. 2 (1994) 267–278.
[7] R.F. Engle, Autoregressive conditional heteroscedasticity with estimator of the variance of United Kingdom inflation, Econometrica 50 (4) (1982) 987–1008.
[8] M. Friedman, The use of ranks to avoid the assumption of normality implicit in the analysis of variance, J. Am. Stat. Assoc. (1937) 675–701.
[9] E. Hadavandi, H. Shavandi, A. Ghanbari, S. Abbasian-Naghneh, Developing a hybrid artificial intelligence model for outpatient visits forecasting in hospitals, Appl. Soft Comput. 12 (2) (2012) 700–711.
[10] W.C. Hong, Y. Dong, L.Y. Chen, S.Y. Wei, SVR with hybrid chaotic genetic algorithms for tourism demand forecasting, Appl. Soft Comput. 11 (2) (2011) 1881–1890.
[11] K.H. Huarng, Effective lengths of intervals to improve forecasting in fuzzy time series, Fuzzy Sets Syst. 123 (2001) 155–162.
[12] S.C. Huang, P.J. Chuang, C.F. Wu, Chaos-based support vector regressions for exchange rate forecasting, Expert Syst. Appl. 37 (12) (2010) 8590–8598.
[13] N.E. Huang, Z. Shen, S.R. Long, A new view of nonlinear water waves: the Hilbert spectrum, Annu. Rev. Fluid Mech. 31 (1999) 417–457.
[14] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis, in: Proceedings of the Royal Society of LONDON Series A: Mathematical Physical and Engineering Sciences, Series A, 454, 1998, pp. 903–995.

[15] K.H. Huarng, T.H.K. Yu, The application of neural networks to forecast fuzzy time series, Phys. A 336 (2006) 481–491.
[16] K.H. Huarng, H.K. Yu, Y.W. Hsu, A multivariate heuristic model for fuzzy, IEEE Trans. Syst. Man Cybern. B: Cybern. 37 (4) (2007) 836–846.
[17] J.S. Jang, ANFIS: adaptive-network-based fuzzy inference systems, IEEE Trans. Syst. Man Cybern. 23 (3) (1993) 665–685.
[18] K.J. Kim, Financial time series forecasting using support vector machines, Neurocomputing 55 (2003) 307–319.
[19] K. Kim, I. Han, Genetic algorithms approach to feature discretization in artificial neural networks for prediction of stock index, Expert Syst. Appl. 19 (2000) 125–132.
[20] T. Kimoto, K. Asakawa, M. Yoda, M. Takeoka, Stock market prediction system with modular neural network, in: Proceedings of The International Joint Conference on Neural Networks, San Diego, California, 1990, pp. 1–6.
[21] L.A. Laboissiere, R.A.S. Fernandes, G.G. Lage, Maximum and minimum stock price forecasting of Brazilian power distribution companies based on artificial neural networks, Appl. Soft Comput. 35 (2015) 66–74.
[22] C.J. Lu, T.S. Lee, C.C. Chiu, Financial time series forecasting using independent component analysis and support vector machine, Decis. Support Syst. 47 (2) (2009) 115–125.
[23] B. Majhi, C.M. Anish, Multiobjective optimization based adaptive models with fuzzy decision making for stock market forecasting, Neurocomputing 167 (2015) 502–511.
[24] C. Narendra Babu, B. Eswara Reddy, A moving-average filter based hybrid ARIMA-ANN model for forecasting time series data, Appl. Soft Comput. 23 (2014) 27–38.
[25] C. Nikolopoulos, P. Fellrath, A hybrid expert system for investment advising, Expert Syst. 11 (4) (1994) 245–250.
[26] D.X. Niu, H.F. Shi, D.D. Wu, Short-term load forecasting using Bayesian neural networks learned by Hybrid Monte Carlo algorithm, Appl. Soft Comput. 12 (6) (2012) 1822–1827.
[27] T.H. Roh, Forecasting the volatility of stock price index, Expert Syst. Appl. 33 (2007) 916–922.
[28] A. Selakov, D. Cvijetinović, L. Milović, S. Mellon, D. Bekut, Hybrid PSO-SVM method for short-term load forecasting during periods with significant temperature variations in city of Burbank, Appl. Soft Comput. 16 (2014) 80–88.
[29] Q. Song, B.S. Chissom, Forecasting enrollments with fuzzy time-series Part I, Fuzzy Sets Syst. 54 (1993) 1–10.
[30] T. Takagi, M. Sugeno, Derivation of fuzzy control rules from human operator's control actions, in: Proceedings IFAC Symposium Fuzzy Information, Knowledge Representation and Decision Analysis, 1983, pp. 55–60.
[31] S. Thawornwong, D. Enke, The adaptive selection of financial and economic variables for use with artificial neural networks, Neurocomputing 56 (2004) 205–232.
[32] C.H.F. Toro, S.G. Meire, J.F. Gálvez, F. Fdez-Riverola, A hybrid artificial intelligence model for river flow forecasting, Appl. Soft Comput. 13 (8) (2013) 3449–3458.
[33] V. Vapnik, The Nature of Statistical Learning Theory, Springer-Verlag, New York, 1995.
[34] H.T. Vincent, S.-L.J. Hu, Z. Hou, Damage detection using empirical mode decomposition method and a comparison with wavelet analysis, in: Proceedings of the Second International Workshop on Structural Health Monitoring, Stanford, 1999, pp. 891–900.
[35] H.K. Yu, Weighted fuzzy time-series models for TAIEX forecasting, Phys. A 349 (2005) 609–624.
[36] D.J. Yu, J.S. Cheng, Y. Yang, Application of EMD method and Hilbert spectrum to the fault diagnosis of roller bearings, Mech. Syst. Signal Process. 19 (2) (2005) 259–270.

Please cite this article in press as: L.-Y. Wei, A hybrid ANFIS model based on empirical mode decomposition for stock time series forecasting, Appl. Soft Comput. J. (2016), http://dx.doi.org/10.1016/j.asoc.2016.01.027