Economic Modelling 30 (2013) 442–448

Contents lists available at SciVerse ScienceDirect

Economic Modelling

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

OWA-based ANFIS model for TAIEX forecasting

Ching-Hsue Cheng [a], Liang-Ying Wei [b,*], Jing-Wei Liu [c], Tai-Liang Chen [d]

[a] Department of Information Management, National Yunlin University of Science and Technology, 123, section 3, University Road, Touliu, Yunlin 640, Taiwan
[b] Department of Information Management, Yuanpei University of Science and Technology, Taipei, Taiwan, No. 306, Yuanpei St., HsinChu 30015, Taiwan
[c] Department of Mutimedia and Game Science, Taipei College of Maritime Technology, No. 212, Sec. 9, Yen Ping N. Rd., Taipei, Taiwan
[d] Department of Information Management and Communication, Wenzao Ursuline College of Languages, 900, Min-Tzu 1st Road, Kaohsiung 807, Taiwan

ARTICLE INFO

Article history:
Accepted 28 September 2012

Keywords:
OWA operator
TAIEX forecasting
ANFIS

ABSTRACT

In stock market forecasting, high-order time-series models that use previous several periods of stock prices as forecast factors are more reasonable to provide a superior investment portfolio for investors than one-order time-series models using one previous period of stock prices. However, in forecasting processes, it is difficult to deal with high-order stock data, because it is hard to give a proper weight to each period of past stock price, reduce data dimensions without losing stock information, and produce a comprehensive forecasting result based on stock data with nonlinear relationships.

Additionally, there are two more drawbacks to past time-series models: (1) some assumptions (Bollerslev, 1986; Engle, 1982) about stock variables are required for statistical methods, such as the autoregressive model (AR) and autoregressive moving average (ARMA); (2) numeric time-series models have been presented to deal with forecasting problems for stock markets, but they can not handle the nonlinear relationships within the stock prices.

To address these shortcomings, this paper proposes a new time series model, which employs the ordered weighted averaging (OWA) operator to fuse high-order data into the aggregated values of single attributes, a fusion adaptive network-based fuzzy inference system (ANFIS) procedure, for forecasting stock price in Taiwanese stock markets.

In verification, this paper employs a seven-year period of the TAIEX stock index, from 1997 to 2003, as experimental datasets and the root mean square error (RMSE) as evaluation criterion. The experimental results indicate that the proposed model is superior to the listing methods in terms of root mean squared error.

1. Introduction

Financial analysts and stock fund managers attempt to predict price activity in the stock market on the basis of either their professional knowledge or with the assistance of stock analyzing tools. If more accurate predictions are given, myriad profit will be made. Therefore, stock analysts have, perennially, strived to discover ways to predict stock price accurately. However, forecasting stock returns are difficult because market volatility needs to be captured in a used and implemented model. Accurate modeling requires, among other factors, consideration of phenomena that is characterized. Stock fund managers and financial analysts predict stock price with their professional knowledge and stock-analyzing tools based on technical analysis (Chuang et al., 2009), fundamental analysis (Atsalakis and Valavanis, 2009), or time series models (Box and Jenkins, 1976; Maia et al., 2008). The area of financial forecasting is very popular for common investors and academy researchers, because for investors, more profit will be made if more accurate predictions are given and for researchers, a higher contribution to research is provided if more proper models are proposed for highly complex stock markets. Therefore, a lot of concepts and techniques are established in fundamental and technical analysis.

Time series models, autoregressive conditional heteroscedastic (ARCH) models (Engle, 1982), Generalized ARCH (GARCH) (Bollerslev, 1986), and autoregressive moving average (ARMA) models (Box and Jenkins, 1976) have been widely utilized to deal with financial forecasting problems in stock markets. However, traditional time series require more historical data along with some assumptions, like normality postulates (Jilani and Burney, 2008).

In the latest decade, the fuzzy set theory proposed by Zadeh (1965) has applied in many forecasting models. In 1993 Song and Chissom (1993a,b), proposed the original model of fuzzy time-series to forecast the enrollments of the University of Alabama. In sequential research, many different fuzzy time series models have been proposed. Chen (1996) proposed another model, which uses equal interval lengths to partition the universe of discourse and generate forecasting rules with a simplified calculation process. Because the length of intervals for the universe of discourse can affect forecasting accuracy, Huarng (2001) proposed distribution-based and average-based lengths to approach

* Corresponding author.
E-mail addresses: chcheng@yuntech.edu.tw (C.-H. Cheng), lywei@mail.ypu.edu.tw (L.-Y. Wei), g9523802@yuntech.edu.tw (J.-W. Liu), 97007@mail.wtuc.edu.tw (T.-L. Chen).

0264-9993/$ – see front matter
http://dx.doi.org/10.1016/j.econmod.2012.09.047

this issue. In the following research, to enhance forecasting performance in stock markets, many fuzzy time-series models employ different weighted technologies to estimate the "recurrence" of stock price patterns, such as a weighted fuzzy time-series method proposed by Yu (2005), and the two weighted fuzzy time-series methods, trend-weighting method (Cheng et al., 2006), and weighted transitional matrix method by Cheng et al. (2008).

Further, to improve the forecasting processes in fuzzy time series with advanced algorithms, Chen and Chung (2006a,b) proposed a new fuzzy time series model that used genetic algorithms to tune the length of linguistic intervals. Additionally, Huarng and Yu (2006) proposed a fuzzy time-series model, which employed a neural network to extract nonlinear fuzzy logical relationships from datasets. However, in order to improve forecasting accuracy, more than one variable in stock markets are considered by financial analysts to build forecasting models. From the research of Huarng et al. (2007), a fuzzy time series model was proposed, which uses the volatility of NASDAQ (the largest US electronic stock market) and Dow Jones (Dow Jones Industrial Average) as forecasting variables to predict the Taiwan stock index. Besides, Cheng and Wei (2009) proposed a volatility model based on a multi-stock index for TAIEX forecasting.

After reviewing the literature of fuzzy time-series models, there are three major drawbacks, summarized as follows. (1) Some assumptions are necessary to build statistical models, such as ARMA, ARCH, and GARCH (Box, and Jenkins,1976; Engle, 1982; Bollerslev, 1986), and therefore they cannot be applied to deal with the datasets with nonlinear relationships (Jilani, and Burney, 2008); (2) past forecasting models can not easily process databases with multiple dimensions, because computation complexity will increase much faster than the growth of data dimension; and (3) the rules produced from the time series models using AI algorithms, such as ANN and GA, are not easily understandable (Chen et al., 2008).

Based on the issues discussed above, the goal of this study is to propose a new forecasting model, which incorporates OWA and ANFIS methods in forecasting processes, to improve the disadvantages of past time series models. OWA is utilized to deal with high dimensions of datasets to reduce computational complexity, and ANFIS is used to build a forecasting model to produce understandable rules for common investors. To verify the proposed model, this paper has provided an empirical analysis using the stock datasets, collected from the Taiwan stock market, in the verification section.

To introduce this study in detail, the rest of the paper is organized as follows. Section 2 describes the related studies. Section 3 briefly presents the proposed model. Section 4 describes the experiments and comparisons. Section 5 presents the findings, and the conclusions of the study.

2. Related works

This section reviews related works of ordered weighted averaging (OWA), the adaptive network-based fuzzy inference system (ANFIS), subtractive clustering (Subclust), and fuzzy time-series.

2.1. Ordered weighted averaging (OWA)

The OWA, which was introduced by Yager, has attracted much interest among researchers (Yager, 1988). Many related studies have been conducted in recent years. For example, Fuller and Majlender (2001) use Lagrange multipliers to derive a polynomial equation to solve constrained optimization problems and to determine the optimal weighting vector.

2.1.1. Yager's OWA

Yager, 1988 proposes an OWA that has the ability to get optimal weights of the attributes based on the rank of these weighting vectors after processing aggregation.

An OWA of dimension n is a mapping f: $R^n \rightarrow R$ that has an associated weighting vector $W=[w_1, w_2, \cdots w_n]^T$ with the following properties:

$$w_i \in [0,1] \quad \text{for} \quad i \in I = \{1,2,\ldots,n\} \quad \text{and} \quad \sum_{i\in I} w_i = 1 \quad \text{such that} \quad f(a_1, a_2, \ldots, a_n) = \sum_{i\in I} w_i b_i \tag{1}$$

where b_i is the ith largest element in the collection of the aggregated objects $\{ a_{1,\ldots,}a_n \}$. Thus, it satisfies the relation $\text{Min}_i[a_i] \le f(a_1,a_2,\ldots,a_n) \le \text{Max}_i[a_i]$ (Ahn, 2006).

Yager (1988) introduced two important characterizing measures with respect to the weighting vector W of an OWA. One of these two measures is *orness* of the aggregation (Chang and Cheng, 2006), which is defined as:

$$orness(W) = (1/n-1)\sum_{i=1}^{n}((n-i) * w_i). \tag{2}$$

where the $orness(W) = \alpha$ is a situation parameter.

The second one is a measure of *dispersion* of the aggregation, which is defined as :

$$\text{Disp(W)} = -\sum_{i=1}^{n} w_i \ln w_i. \tag{3}$$

The *orness* measure has the following property:

$W = \{w_1, w_2 \ldots, w_n\}$ is the weight vector of an OWA with $orness(W) = \alpha$. Then,
$W' = \{w_n, w_{n-1}, \ldots, w_1\}$ is the reverse order of W, $orness(W') = 1-\alpha$.

O'Hagan (1988) combines the principle of maximum entropy and OWA to propose a particular OWA that has maximum entropy with a given level of *orness*. The definition is presented as follows:

$$\text{Maximize the function} \quad -\sum_{i=1}^{n} w_i \ln w_i. \qquad \alpha = \frac{1}{n-1}\sum_{i=1}^{n}(n-i)w_i, \quad 0 \le \alpha \le 1. \tag{4}$$

2.1.2. Fuller and Majlender's OWA

Fuller and Majlender (2001) transform Yager's OWA equation to a polynomial equation by using Lagrange multipliers. According to their approach, the associated weighting vector can be obtained by Eq. (5)–(7)

$$\ln w_j = \frac{j-1}{n-1}\ln w_n + \frac{n-j}{n-1}\ln w_1 \Rightarrow w_j = \sqrt[n-1]{w_1^{n-j} w_n^{j-1}}. \tag{5}$$

$$\text{and} \quad w_1[(n-1)\alpha + 1 - nw_1]^n = [(n-1)\alpha]^{n-1}[((n-1)\alpha - n)w_1 + 1]. \tag{6}$$

$$\text{if} \quad w_1 = w_2 = \ldots\ldots = w_n = \frac{1}{n} \Rightarrow \text{disp}(W) = \ln n \quad (\alpha = 0.5) \qquad \text{then} \quad w_n = \frac{((n-1)\alpha - n)w_1 + 1}{(n-1)\alpha + 1 - nw_1}. \tag{7}$$

where w_i is the weight vector, n is the number of attributes, and α is the situation parameter.

2.2. Subtractive clustering

Chiu (1994) developed subtractive clustering, a type of fuzzy clustering, to estimate both the number and initial locations of cluster centers. Consider a set T of N data points in a D-dimensional hyperspace, where each data point W_i ($i = 1,2,\ldots,N$). $W_i = (x_i, y_i)$, 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. (8)

$$P_i = \sum_{j=1}^{N} e^{-\alpha\|W_i - W_j\|^2}, \tag{8}$$

where $\alpha = {}^4/_{r^2}$, r is the radius defining a W_i neighborhood, and ||.|| denotes the Euclidean distance.

A data point with many neighboring data points will have a high potential value, and the data point with the highest potential value 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 by Eq. (9)

$$p_i = p_i - p_1^* \exp\left(-\beta\left\|W_i - W_1^*\right\|^2\right), \tag{9}$$

where β is a positive constant, defining 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. (9), the method selects the data point with the highest remaining potential as the second cluster center. For the general equation, we can rewrite Eq. (9) as Eq. (10).

$$p_i = p_i - p_k^* \exp\left(-\beta\left\|W_i - W_k^*\right\|^2\right), \tag{10}$$

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.

2.3. ANFIS: adaptive network-based fuzzy inference system

Jang (1993) proposed ANFIS, which is a fuzzy inference system implemented in the framework of adaptive networks. For illustrating the system, we assume that the fuzzy inference system consists of five layers of adaptive network with two inputs x and y and one output z. The architecture of ANFIS is shown in Fig. 1.

Then, we suppose that the system consists of 2 fuzzy if-then rules based on Takagi and Sugeno's type (1983):

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. (11)). $O_{1,i}$ is the membership function of A_i, and it specifies the degree to which the given x satisfies the quantifier A_i. Usually, we select the bell-shaped membership function as the input membership function (see Eq. (12)), with maximum equal to 1 and minimum equal to 0.

$$O_{1,i} = \mu A_i(x) \quad \text{for}\, i = 1, 2, \tag{11}$$

$$\mu A_i(x) = \frac{1}{1 + \left[\left(\frac{x - c_i}{a_i}\right)^2\right]^{bi}}, \tag{12}$$

where a_i, and b_i vary the width of the curve, b_i is a positive value, and c_i 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 by Eq. (13).

$$O_{2,i} = w_i = \mu A_i(x) \times \mu B_i(y) \quad \text{for}\, i = 1, 2. \tag{13}$$

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 sum of all rules' firing strengths by Eq. (14). Output of this layer can be called normalized firing strengths.

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad \text{for}\, i = 1, 2. \tag{14}$$

Layer 4: Every node i in this layer is a square node with a node function (see Eq. (15)). Parameters in this layer will be referred to as consequent parameters.

$$O_{4,i} = \bar{w}_i f_i = \bar{w}_i (p_i + q_i + r_i), \tag{15}$$

where p_i, q_i, and r_i are the parameters.

Layer 5: The single node in this layer is a circle node labeled $\sum$ that computes the overall output as the summation of all incoming signals (see Eq. (16))

$$O_{5,i} = \sum_i \bar{w}_i f_i = \frac{\sum_{i=1} w_i f}{\sum_{i=1} w_i} = \text{overall output}. \tag{16}$$

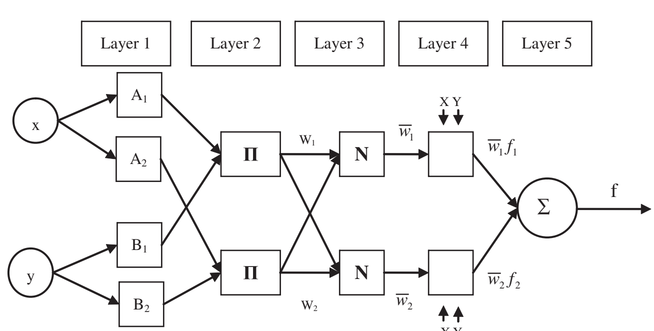

Fig. 1. The architecture of the ANFIS network.

2.4. Fuzzy time-series

Song and Chissom, 1993a,b proposed a fuzzy time series model to deal with the problems involving human linguistic terms (Zadeh, 1975a,b; Zadeh, 1976; Ross, 1995). In the following research, they continued to discuss the difference between time-invariant and time-variant models (Song and Chissom, 1993a,b; Song, and Chissom, 1994). Besides the above researchers, Chen (1996) presented a method to forecast the enrollments of the University of Alabama based on fuzzy time series.

In the past years, many fuzzy time-series models have been applied by following Song and Chissom's definitions (Bollerslev, 1986; Song and Chissom, 1993a,b; Song, and Chissom, 1994). Among these models, Chen's model is the very conventional one because of easy computations and good forecasting performance (Chen, 1996). The definitions and processes of the fuzzy time-series presented by Song and Chissom (1993a,b) are described as follows:

Definition 1. fuzzy time-series.

Let Y(t) (t = …,0,1,2,…), a subset of real numbers, be the universe of discourse by which fuzzy sets $f_j(t)$ are defined. If $F(t)$ is a collection of $f_1(t)$, $f_2(t)$…then $F(t)$ is called a fuzzy time-series defined on y(t).

Definition 2. fuzzy time-series relationships.

Assuming that $F(t)$ is caused only by $F(t-1)$, then the relationship can be expressed as: $F(t)=F(t-1)*R(t, t-1)$, which is the fuzzy relationship between $F(t)$ and $F(t-1)$, where $*$ represents an operator. To sum up, let $F(t-1)=$ Ai and $F(t)=$ Ai. The fuzzy logical relationship between $F(t)$ and $F(t-1)$ can be denoted as $A_i \rightarrow A_j$, where Ai refers to the left-hand side and A_j refers to the right-hand side of the fuzzy logical relationship. Furthermore, these fuzzy logical relationships can be grouped to establish different fuzzy relationships.

3. Proposed model

3.1. The proposed concepts

From literature reviews in Section 1, there are three major drawbacks, found as follows:

Firstly, some assumptions, such as stationarity, and non-stationarity, are necessary to build statistical models, such as ARMA, ARCH, and GARCH (Box and Jenkins, 1976; Engle, 1982; Bollerslev, 1986). In stock markets, the relationships between past stock prices and the future prices are not only linear but also nonlinear, and therefore the statistical models can not build proper models for stock datasets with linear and nonlinear relationships (Jilani and Burney, 2008).

Secondly, stock datasets usually consist of many different variables (technique indicators) or multiple periods of stock prices, and therefore forecasting models have to deal with a high dimension of datasets to produce forecasting results. It is critical to fuse variables into useful forecasting factors without losing information contained in the original variables to produce predictions, and it is a common situation for forecasting models that there is a tradeoff between data dimension and data information when fusing forecasting variables into fewer factors. Besides, for forecast models using multiple variables, computational complexity will increase much faster than the growth of data dimension.

Finally, in the recent research, ANN and GA have widely been applied in fuzzy time-series (Chen and Chung, 2006a,b; Huarng and Yu, 2006; Zhang, 2003) and the fuzzy logic relationships (FLRs) are extracted by these algorithms to build forecasting models. However, the mining process is not easily understood, like the black box, and the forecasting rules are not useful for investors to buy and sell stocks.

ANFIS can process data without any assumption about datasets and discover nonlinear relationships between observations, and therefore we argue that it is proper to overcome the drawbacks, mentioned as follows. (1) Statistical models, such as ARMA, ARCH, and GARCH (Bollerslev, 1986; Box and Jenkins, 1976; Engle, 1982), need some assumptions (stationarity and non-stationarity Bollerslev, 1986; Engle, 1982) to build models for stock markets; and (2) statistical methods only can deal with linear relationships, but linear and nonlinear relationships exist between past stock prices and the future prices in stock markets. Furthermore, ANFIS can produce “if-then” rules (see Section 2.3) to model the qualitative aspects of human knowledge, and the rules from ANFIS are understandable and applicable for investors.

To overcome the fact found in the past research for stock forecasting, the OWA operator is proposed to reduce the data dimension of stock datasets, and a liner equation is employed to fuse recent periods of prices with the given weight into one aggregate value as input for the ANFIS model to produce the optimal forecast for the future stock price. OWA utilized a simple method to reduce data dimension with one linear equation. In the OWA fusion process, each attribute is multiplied with its corresponding weight, and the weighted attributes are summarized as one aggregated value. Therefore, data dimension can be reduced with less computation because of the linear calculation of OWA.

Besides, in practical stock markets, investors usually make their short-term decisions based on recent stock information, such as recent periods of stock prices and technical analysis reports. Therefore, the investors are influenced by recent stock price fluctuations to make investment decisions, and the recent different periods of prices influence their decisions to different degrees. To deal with the condition above, the OWA operator is proposed to produce one aggregate value for forecasting the future prices with different weights to meet the variations of influence degrees. Additionally, we argue that applying OWA in forecasting models can improve forecasting accuracy, because the literature (Fuller and Majlender, 2001) has shown that weighted-based technologies are superior methods to enhance the performances of time series models.

Based on concepts mentioned above, this paper proposes a novel model, which combines OWA (OWA operator is proposed to produce one aggregate value for forecasting the future prices with different weights to meet the variations of influence degrees) and ANFIS method (ANFIS can produce “if-then” rules to model the qualitative aspects of human knowledge) performance, and the three procedures are provided in the proposed model, as follows (the overall flowchart of the proposed model is shown in Fig. 2).

In the first procedure, use OWA operators based on several α values (which means different conditions in the stock market), from 0.5 to 1.0, to produce different weights (w1, w2, and w3) for recent periods of stock prices (TAIEX (t), TAIEX ($t-1$), TAIEX ($t-2$)) by Eq. (5) to Eq. (7). The maximum weight value (w1) is assigned to TAIEX (t), the middleweight value (w2) is assigned to TAIEX ($t-1$), and the minimum weight value (w3) is assigned to TAIEX ($t-2$).

In the second procedure, the recent three periods of stock prices, TAIEX (t), TAIEX ($t-1$), TAIEX ($t-2$), are integrated into one value with their weights by one linear equation.

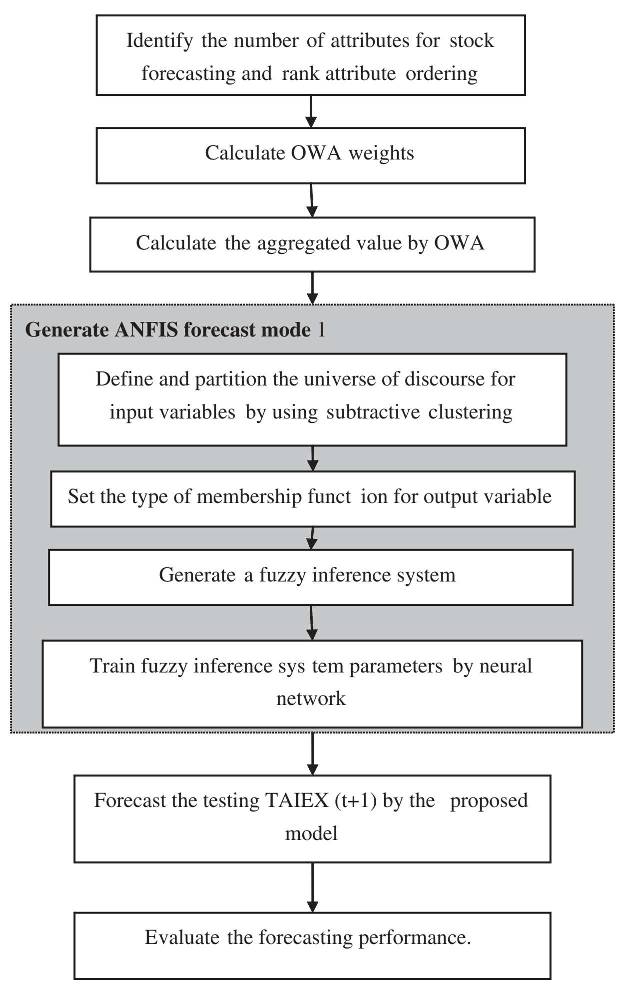

Fig. 2. Flowchart of proposed procedure.

In the third procedure, the ANFIS model is employed to produce the final forecast for the future stock price and uses the integrated value from the second procedure as the input value for ANFIS model. In this procedure, the ANFIS model can optimize fuzzy inference system parameters by the adaptive network, which can overcome the limitations of statistical methods (data need to obey some mathematical distribution).

To actualize the proposed concepts, an algorithm is proposed in this paper, and introduced in the following section.

3.2. The proposed algorithm

The algorithm of the proposed method is introduced step by step, as follows.

Step 1. Identify the amount of stock price periods (number of attribute) for forecasting the future price.

The research advanced by Chen et al. (2008) has revealed that the price patterns in the TAIEX are short-term. In this step, the recent three periods of stock prices, P(t−2), P(t−1), and P(t), are utilized to forecast the future price, P(t+1), and identify the number of attributes as three. In a practical stock market, we argue that the influence degree of past stock price would decrease gradually by a higher order of lag periods. Therefore, in this step the ordering is set as follows: P(t) > P(t−1) > P(t−2).

Step 2. Calculate the OWA weights and assign the influence degree for the recent stock price.

In this step, a refined OWA algorithm (Cheng et al., 2009) is provided to calculate the OWA weights by Eqs. (5)–(7), and the main process of this algorithm is presented as Fig. 3.

From the OWA algorithm, different α values, from 0.5 to 1.0, are used to produce different sets of OWA weights. In this step, each α value represents one set of influence degree, which consists of W_1, W_2, and W_3, for the recent three periods of stock prices (see Table 1).

Step 3. Calculate the aggregated value.

In this step, with influence degree and stock price, we calculate the aggregated value by one linear equation, defined in Eq. (9).

$$A(t) = W_1 \times P(t) + W_2 \times P(t-1) + W_3 \times P(t-2) \tag{9}$$

where W_1, W_2, and W_3 denote different influence degrees for the three recent periods of stock prices, P(t), P(t−1), and P(t−2), and A(t)

OWA (n, α)

/* n is the number of attribute; α is the situation parameter */

If $\alpha < 0.5$

Then $\alpha = 1 - \alpha$

If $\alpha > 0.5$

Then $w_1[(n-1)\alpha+1-nw_1]^n = [(n-1)\alpha]^{n-1}[((n-1)\alpha-n)w_1+1]$ //Calculate

$w_n = [((n-1)\alpha-n)w_1+1]/[(n-1)\alpha+1-nw_1]$ //Calculate w_n

For i = 2 to n-1 do

$w_i = \sqrt[n-1]{w_1^{n-i} w_n^{i-1}}$ //Calculate w_i

Fig. 3. The algorithm of OWA [14].

Table 1
OWA weights when n = 3

	α = 0.5	α = 0.6	α = 0.7	α = 0.8	α = 0.9	α = 1.0
W_1	0.3333	0.4384	0.5540	0.6819	0.8263	1
W_2	0.3333	0.3232	0.2920	0.2358	0.1470	0
W_3	0.3333	0.2384	0.1540	0.0819	0.0263	0

represents an aggregated value of the three recent periods of stock prices with their corresponding influence degrees.

Step 4. Generate ANFIS forecast model.

In this step, the ANFIS model with the subtractive clustering method is employed to produce forecasting rules and a final forecast is generated for the future price. There are two sub-steps included in this step and described as follows:

Step 4.1. Define and partition the universe of discourse for input variables by the subtractive clustering method.

Firstly, we define the universe of discourse for input variables (A(t), an aggregated value of the three recent periods of stock prices) according to its minimum and maximum value. Secondly, partition the universe of discourse to produce linguistic intervals by subtractive clustering (Chiu, 1994) (Gaussian membership function).

Step 4.2. Set the type of membership function for output variables.

In this step, one type membership function is set to produce output variables. From step 4.1, there is only one aggregated value, A(t), which is the aggregated value of P(t), P (t−1) and P(t−2), used as one input variable. To produce linguistic variables as output variables for the next step, one type membership function is employed based on the three linguistic intervals produced by the subtractive clustering method in step 4.1.

Step 4.3. Generate a fuzzy inference system.

In this step, the Sugeno fuzzy model 1983 is employed to generate fuzzy if-then rules, and a typical rule in the fuzzy inference system is described as follows:

$$\text{If } x(A(t)) = A_i \quad \text{then} \quad f_i(P(t+1)) = p_i x + r_i$$

where x (A(t)), is linguistic variables, A_i is the linguistic values (high, middle, low), f_i denotes the i-th output value, p_i and r_i, are the parameters (i = 1, 2, 3), the linguistic values (A_i) from input membership functions are used as the if-condition part, and the output membership functions (f_i) is the then part.

Step 4.4. Train fuzzy inference system parameters by neuron network.

In this section, we employ a combination of the least-squares method and the backpropagation gradient-descent method for training four types of forecasting models and use fuzzy inference system membership function parameters to emulate a given training dataset. This study sets the epoch as 50 (the process is executed for the predetermined fixed number (50) of iterations unless it terminates while the training error converges) for the training stopping criterion and then obtains the parameters for the selected output membership function.

In the proposed algorithms, a_j is the aggregated value of TAIEX (T), TAIEX (T−1) and TAIEX (T−2), and the value is used as forecasting factor to forecast the future stock index, TAIEX (t+1). There are three linguistic values (low, middle, and high) produced from the subtractive clustering method and, therefore, three decision rules (the amount of rules is equal to the amount of linguistic values) are generated and described as follows:

Rule 1:

$$\text{If}\, x\left(a_j\right) = A_{low} \quad \text{then} \quad f_{low}(\text{TAIEX}(t+1)) = p_{low}x + r_{low}$$

Rule 2:

$$\text{If}\, x\left(a_j\right) = A_{middle} \quad \text{then} \quad f_{middle}(\text{TAIEX}(t+1)) = p_{middle}x + r_{middle}$$

Rule 3:

$$\text{If}\, x\left(a_j\right) = A_{high} \quad \text{then} \quad f_{high}(\text{TAIEX}(t+1)) = p_{high}x + r_{high},$$

where x (a_j) is the linguistic variables, A_i is the linguistic values, $f_i(\text{TAIEX}(t+1))$ denotes the *i-th* output value, and p_i, r_i, are the parameters ($i=1, 2, 3$ denotes low, middle, high).

For example, we take rule 1 ($\alpha=1$) generated by training data in 1998 as an explanation.

If a_j (training data in 1998) is low, then TAIEX $(t+1)=0.5285\times a_j+3715$.

Step 5. Forecast the P $(t+1)$ by the proposed model.

The parameters of the proposed model (p_i, r_i, mentioned in step 5.3) are determined when the stopping criterion is reached (from step 5.4), and then employ the optimal forecasting model in training datasets to forecast the future price $P(t+1)$ in the target testing datasets.

Step 6. Evaluate the forecasting performance.

In this step, the RMSE, defined by Eq. (15), is taken as an evaluation criterion to compare with other forecasting models.

$$\sqrt{\sum_{t=1}^{n} |\text{actual}(t) - \text{forecast}(t)|^2 \Big/ n}. \tag{15}$$

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

4. Model verification

This section has provided an evaluation for the proposed model in forecasting performance and model comparisons using the two recent fuzzy time series models (Chen, 1996; Huarng and Yu, 2006). The empirical analysis in this paper selects a seven-year period of TAIEX stock data, from 1997 to 2003, as experimentation datasets. Each year of TAIEX is used as one complete verification dataset, the first 10-month period of each year is used as a training dataset, and the rest, from November to December, is selected as a testing dataset. Besides, this paper employs the RMSE (see Eq. (15)) as a performance indicator, and three forecasting models, (Chen, 1996; Huarng and Yu, 2006), and NN-based fuzzy time-series, as comparison models.

4.1. Experimental results

In the experiments, there are 42 forecasting performance records (see Table 2) for the proposed model using different α values (from 0.5 to 1, stepped value is 0.1) for the 7 testing datasets. The performance comparisons for the proposed model, Chen's model, and Huarng's model are listed in Table 2 and illustrated in Fig. 4. From Table 2, it has been shown that the proposed model outperforms the listing comparison models in the 7 testing datasets under the α values, from 0.5 to 1, and the average performance of the proposed model is also superior to the comparison models.

For making some simulation trades and showing the profits, we set two trade rules by the Taiwan Futures Exchange (TAIFEX) and use the two trade rules to calculate profits, and assume that the profit unit is equal to one. So the profit formula is defined as Eq. (16).

Rule 1: sell rule

$$\text{If}\ \frac{|\text{forecast}(t) - \text{actual}(t)|}{\text{actual}(t)} \le \beta \quad \text{and} \quad \text{forecast}(t+1) - \text{actual}(t) > 0 \text{ then sell}$$

Rule 2: buy rule

$$\text{If}\ \frac{|\text{forecast}(t) - \text{actual}(t)|}{\text{actual}(t)} \le \beta \quad \text{and} \quad \text{forecast}(t+1) - \text{actual}(t) < 0 \text{ then buy}$$

where β denotes threshold parameter ($0<\beta\le 0.07$, the threshold parameter depends on daily fluctuation of TAIEX).

Profit definition.

$$\text{Profit} = \sum_{t_s=1}^{p} (\text{actual}(t+1) - \text{actual}(t)) + \sum_{t_b=1}^{q} (\text{actual}(t) - \text{actual}(t+1)) \tag{16}$$

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 β is obtained when the forecasting performance reaches best profits in the training dataset. From the optimal threshold parameter β and Eq. (16), the profits for different models are calculated and the profit results are shown in Table 3. From the profits comparison table (see Table 3), we can see that the proposed model has higher profits than the listing models.

5. Findings and conclusions

This paper has proposed a new time series model, which has fused two AI methods, the OWA and ANFIS model, in forecasting processes to enhance prediction accuracy. In the proposed model, the OWA operator is employed to fuse high-order data into one aggregated value, and the ANFIS model is utilized for forecasting stock price with the aggregated value produced by OWA. From the experimental results, it is discovered that the proposed model outperforms the listing methods in forecasting accuracy, and there are three findings, provided as follows.

Moreover, the results of this paper are useful and viable for stock investors, decision makers and future researches. Investors can utilize

Table 2
The performance comparisons of different models (TAIEX).

Models	Dataset	Year						
		1997	1998	1999	2000	2001	2002	2003
Chen's model (1996)		154	134	120	176	148	101	74
Huarng and Yu's model (2006)		141	121	109	152	130	84	56
Proposed model	($\alpha=0.5$)	152	146	142	201	168	81	69
	($\alpha=0.6$)	145	136	132	183	156	76	65
	($\alpha=0.7$)	139	128	122	166	144	72	62
	($\alpha=0.8$)	135	121	113	154	134	69	59
	($\alpha=0.9$)	133[a]	117	107	133	125	66[a]	56
	($\alpha=1.0$)	133[a]	115[a]	103[a]	130[a]	120[a]	66[a]	55[a]

[a] The best performance among the listing models.

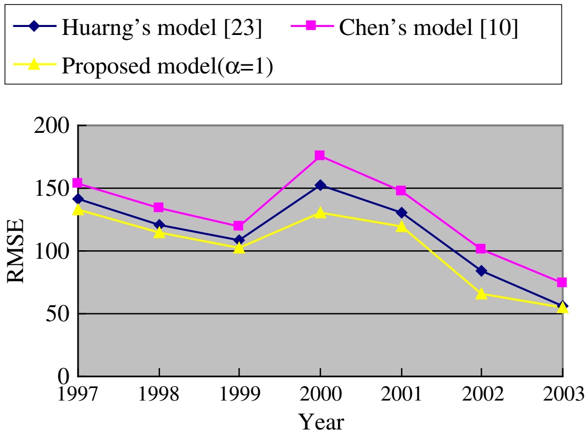

Fig. 4. The performance comparisons of different models (TAIEX).

this "if-then" rule generated from ANFIS discover the superior target of investment with benefits in stock market. Further, the more reasonable and understandable rules produced by ANFIS can model the qualitative aspects of human knowledge.

(1) The OWA operator is an important tool to cope with multi-attribute data, and it is different from classical weighted methods. Further, the OWA operator can adjust the weight of attributes based on the situation of the decision-maker and aggregate different attribute values into a single aggregated value of each instance, and then the single aggregated value is utilized to generate fuzzy if-then rules by the ANFIS forecast model for forecasting performance. According to Table 2 and Fig. 4, it shows that the proposed model is superior to the listing methods in terms of RMSE.
(2) From the empirical results, the best performance of the proposed method generated is $\alpha = 1$; that is, the weights for TAIEX (t), TAIEX ($t-1$), and TAIEX ($t-2$) are 1, 0, and 0, respectively. This means that the last stock index TAIEX (t) significantly affects TAIEX in the next day, and Chen et al.'s paper (2008) shows that stock price patterns in the Taiwan stock market are short term.
(3) Few forecasting rules are generated (only three rules; the number of rules is the same as the number of linguistic intervals by subtractive clustering) to forecast electricity loads.

In the future works, we can use other stock datasets, such as HSI (http://www.hsi.com.hk/HSI-Net/); NASDAQ (http://www.nasdaq.com/), and DJI (http:/www.djindexes.com/), to validate the proposed model further. Moreover, there are two approaches that can be applied to the proposed model that may improve forecasting accuracy: (1) employ other machine learning methods, such as SVM and rough set algorithm, to enhance performance of the proposed model; and (2) combine more effective financial factors, such as technical indicators, with the proposed method to improve accuracy.

Table 3
The profits comparisons of different models (TAIEX).

Year	β	Models		
		Yu's model [25]	Chen's model [4]	Proposed model ($\alpha = 1.0$)
1997	0.02	−107	−127	825[a]

[a] The best profits among three models.

References

Ahn, S.B., 2006. On the properties of OWA operator weights functions with constant level of orness. IEEE Transactions on Fuzzy System 14 (4), 511–515.
Atsalakis, G.S., Valavanis, K.P., 2009. Forecasting stock market short-term trends using a neuro-fuzzy based methodology. Expert Systems with Applications 36 (7), 10696–10707.
Bollerslev, T., 1986. Generalized autoregressive conditional heteroscedasticity. Journal of Econometrics 31, 307–327.
Box, G., Jenkins, G., 1976. Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco.
Chang, J.R., Cheng, C.H., 2006. MCDM aggregation model by ME-OWA and ME-OWGA operators. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 14 (4), 421–443.
Chen, S.M., 1996. Forecasting enrollments based on fuzzy time-series. Fuzzy Sets and Systems 81, 311–319.
Chen, S.-M., Chung, N.-Y., 2006a. Forecasting enrollments of students by using fuzzy time series and genetic algorithms. International Journal of Information and Management Sciences 17, 1–17.
Chen, S.M., Chung, N.Y., 2006b. Forecasting enrollments using high-order fuzzy time series and genetic algorithms. International of Intelligent Systems 21, 485–501.
Chen, T.-L., Cheng, C.-H., Teoh, H.-J., 2008. High-order fuzzy time-series based on multi-period adaptation model for forecasting stock markets. Physica A: Statistical Mechanics and its Applications 387 (4), 876–888.
Cheng, C.-H., Wei, L.-Y., 2009. Volatility model based on multi-stock index for TAIEX forecasting. Expert Systems with Applications 36 (3), 6187–6191.
Cheng, C.H., Chen, T.L., Chiang, C.H., 2006. Trend-weighted fuzzy time-series model for TAIEX forecasting. Lecture Notes in Computer Science 4234, 469–477.
Cheng, C.H., Wang, J.W., Li, C.H., 2008. Forecasting the number of outpatient visits using a new fuzzy time series based on weighted-transitional matrix. Expert Systems with Applications 34, 2568–2575.
Cheng, C.-H., Wang, J.-W., Wu, M.-C., 2009. OWA-weighted based clustering method for classification problem. Expert Systems with Applications 36 (3, Part 1), 4988–4995.
Chiu, S.L., 1994. Fuzzy model identification based on cluster estimation. Journal of Intelligent and Fuzzy Systems 2, 267–278.
Chuang, C.-C., Kuan, C.-M., Lin, H.-Y., 2009. Causality in quantiles and dynamic stock return-volume relations. Journal of Banking & Finance 33 (7), 1351–1360.
Dow Jones Indexes, d. http:/www.djindexes.com/.
Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimator of the variance of United Kingdom inflation. Econometrica 50 (4), 987–1008.
Fuller, R., Majlender, P., 2001. An analytic approach for obtaining maximal entropy OWA operator weights. Fuzzy Sets and Systems 124, 53–57.
Hang Seng Index, d. http://www.hsi.com.hk/HSI-Net/.
Huarng, K.H., 2001. Effective lengths of intervals to improve forecasting in fuzzy time series. Fuzzy Sets and Systems 123, 155–162.
Huarng, K.H., Yu, H.K., 2006. The application of neural networks to forecast fuzzy time series. Physica A 336, 481–491.
Huarng, K.H., Yu, H.K., Hsu, Y.W., 2007. A multivariate heuristic model for fuzzy time-series forecasting. IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics 37 (4), 836–846.
Jang, J.S., 1993. ANFIS: adaptive-network-based fuzzy inference systems. IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics 23 (3), 665–685.
Jilani, T.A., Burney, S.M.A., 2008. A refined fuzzy time series model for stock market forecasting. Physica A 387, 2857–2862.
Maia, A.L.S., de Carvalho, F.d.A.T., Ludermir, T.B., 2008. Forecasting models for interval-valued time series. Neurocomputing 71 (16–18), 3344–3352.
NASDAQ, d. http://www.nasdaq.com/.
O'Hagan, M., 1988. Aggregating Template Rule Antecedents in Real-time Expert Systems with Fuzzy Set Logic. Proc 22nd Annual IEEE Asilomar Conference on Signals, Systems and Computers. IEEE and Maple Press, Pacific Grove, CA, pp. 681–689.
Ross, T.J., 1995. Fuzzy Logic with Engineering Applications. McGraw-Hill, New York.
Song, Q., Chissom, B.S., 1993a. Forecasting enrollments with fuzzy time-series Part I. Fuzzy Sets and Systems 54, 1–10.
Song, Q., Chissom, B.S., 1993b. Fuzzy time series and its models. Fuzzy Sets and Systems 54, 269–277.
Song, Q., Chissom, B.S., 1994. Forecasting enrollments with fuzzy time-series — part II. Fuzzy Sets and Systems 62, 1–8.
Takagi, T., Sugeno, M., 1983. Derivation of fuzzy control rules from human operator's control actions. IFAC Symp. Fuzzy Inform., Knowledge Representation and Decision Analysis, pp. 55–60.
Yager, R.R., 1988. On ordered weighted averaging aggregation operators in multi-criteria decision making. IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics 183–190.
Yu, H.K., 2005. Weighted fuzzy time-series models for TAIEX forecasting. Physica A 349, 609–624.
Zadeh, L., 1965. Fuzzy sets. Information and Control 8, 338–353.
Zadeh, L.A., 1975a. The concept of a linguistic variable and its application to approximate reasoning II. The Information of the Science 8, 301–357.
Zadeh, L.A., 1975b. The concept of a linguistic variable and its application to approximate reasoning, part I. Information Sciences 8, 199–249.
Zadeh, L.A., 1976. The concept of a linguistic variable and its application to approximate reasoning III. The Information of the Science 19, 43–80.
Zhang, G.P., 2003. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing 50, 159–175.