8. A DYNAMIC WEIGHTED DISTANCE-BASED FUZZY TIME SERIES NEURAL NETWORK WITH BOOTSTRAP MODEL FOR OPTION PRICE FORECASTING

Chih-Chung YANG[1]
Yungho LEU[1]
Chien-Pang LEE[2*]

Abstract

The option price forecasting is still a big challenging problem because the option pricing is determined by many factors. Accordingly, it is difficult to predict option price accurately. To counter this problem, this paper proposes a novel hybrid model to forecast the option price. The proposed model, termed as the dynamic weighted distance-based fuzzy time series neural network with bootstrap model, is composed of a dynamic n-order 2-factor fuzzy time series model, a radial basis function neural network model and a bootstrap method. In the proposed model, the dynamic n-order 2-factor fuzzy time series model can automatic choose the best n-order for searching similar data from historical data and, then, build a training dataset for the radial basis function neural network model to forecast the option price. However, the sample size of option price data is small. Accordingly, this paper uses the bootstrap method to enhance the prediction accuracy of the proposed model. The experiment results show that the proposed model outperforms several existing methods in terms of RMSE, MAE and the testing results of Diebold-Marioano test.

Keywords: option price forecasting, fuzzy time series model, radial basis function neural network model, bootstrap method, Diebold-Marioano test

JEL Classification: C32, C53, C63

1 Department of Information Management, National Taiwan University of Science and Technology, Taiwan.
2* Department of Information Management, Da-Yeh University, Taiwan. Corresponding author, E-mail: chien.pang@gmail.com.

I. Introduction

Recently, many investors want to earn their own profits by investing in the financial market. Many investing targets are in the financial market, such as stocks, bonds, futures, and funds. How to choose a suitable investing target is important for investors. However, high profit brings high risk. Accordingly, it is an important research issue to reduce the risk of investments. Due to the fact that the option is one of the important tools for risk management in financial investments (Ko, 2009; Leu *et al*., 2010, 2011), the transaction of the option has become more popular in the recent years (Huang, 2008). For example, a producer can buy a put option to prevent a profit loss due to decrease in the price of his products in the future. Similarly, a customer can buy a call option to buy his desired products at an expected price in the future (Leu *et al*., 2010, 2011). Hence, an option is like an insurance policy in that people have to pay premium for an option. The premium, also called the price, of an option will affect investors to buy this option. However, the price of an option is determined by many factors, such as the current stock price, the option strike price, the time to expiration, the volatility of the stock price and the risk-free interest rates (Black and Scholes, 1973). Consequently, the option price forecasting is still a big challenging problem, because the option price is affected by many factors. The well-known Black-Scholes model (the B-S model) (Black and Scholes, 1973) was first introduced for option pricing in 1973. The B-S model is limited due to the fact that many of its assumptions might not be fitted in the real life financial issue. To solve this problem, many researchers propose novel methods to predict the option price of the real life financial issue in the recent years (Mercuri, 2008).

Many researchers used data mining models to forecast the financial research issues in the recent years. For example, due to the fact that the artificial neural network (ANN) does not need any a priori knowledge of the distribution (Morariu, 2009; Saman, 2011), ANN is the most popular prediction model in financial forecasting, such as the stock indices forecasting (Alhassan and Misra, 2011; Grudnitski and Osburn, 1993; Morelli *et al*., 2004; Takahama *et al*., 2009) and exchange rate forecasting (Panda and Narasimhan, 2007; Shin and Han, 2000). Moreover, the ANN model is also widely applied in option price forecasting research issues. For example, many researchers proposed hybrid neural network models for option price forecasting (Ko, 2009; Lajbcygier, 2004; Lajbcygier and Conner, 1997; Tseng *et al*., 2008; Wang, 2009). In addition, the fuzzy time series model is gaining its popularity in financial forecasting issues in the recent years, such as stock indices forecasting (Cheng *et al*., 2008; Teoh *et al*., 2007; Yu, 2005), foreign exchange rate forecasting (Leu *et al*., 2009) and futures exchange indices forecasting (Lee *et al*., 2006).

In this paper, we propose a novel hybrid model to predict the option price of “Taiwan Stock Exchange Stock Price Index Options (TXO)”. The proposed model, called the dynamic weighted distance-based fuzzy time series neural network with bootstrap model, is composed of a dynamic *n*-order *2*-factor fuzzy time series model, a radial basis function neural network model and a bootstrap method. The main idea of the proposed model is that the dynamic *n*-order *2*-foctor fuzzy time series model can automatically choose the best *n*-order and use the selected *n*-order to search the similar data from historical data to generate a training dataset for the radial basis function neural network model to predict the option price. Finally, due to the fact that the sample size of training data of option price is small, in this paper the bootstrap method is used to enhance the prediction accuracy.

The remainder of this paper is organized as follows. Section 2 reviews the related methods, including the definitions of fuzzy time series model, that of the artificial neural networks model and that of bootstrap method. Section 3 introduces the procedure of the proposed model. Section 4 compares the performance of the proposed model with the other existing methods. Section 5 gives the conclusions of this paper.

2. Related Methods

Due to the fact that the fuzzy time series model, the radial basis function neural network model and the bootstrap method play important roles in this paper, we briefly review the definition of the fuzzy time series model in Section 2.1 and introduce the concept of the radial basis function neural network model in Section 2.2. Finally, Section 2.3 reviews the definition of the bootstrap method.

2.1 The Fuzzy Time Series Model

A fuzzy time series model, which is based on the fuzzy logic, is used to solve problems in forecasting (Wang *et al*., 2009). Song and Chissom (1993a) first applied it for forecasting the enrollments at the University of Alabama. Recently, the fuzzy time series model is also widely used in many financial issues (Leu *et al*., 2009). According to the literature (Chen, 2002; Lee *et al*., 2006; Leu *et al*., 2010, 2011; Song and Chissom, 1993a, b, 1994), the following definitions are given to a fuzzy time series model.

Definition 1:

Let $Y(t)$ (t =…,0,1,2,…), a subset of R, be the universe of discourse in which fuzzy sets $f_i(t)$ (i = 1,2...) are defined. If $F(t)$ is a collection of $f_i(t)$, $F(t)$ is called a fuzzy time series defined on $Y(t)$.

Definition 2:

If for any $f_j(t)\in F(t)$, there is a $f_i(t\text{-}1)\in F(t\text{-}1)$, such that there is a fuzzy relation $R_{ij}(t, t\text{-}1)$ and $f_j(t)=f_i(t\text{-}1)\circ R_{ij}(t, t\text{-}1)$ where '$\circ$' is the max-min composition, $F(t)$ is said to be caused by $F(t\text{-}1)$ only, and it can be represented by $F(t\text{-}1)\rightarrow F(t)$.

Definition 3:

If $F(t)$ is caused by $F(t\text{-}1)$, $F(t\text{-}2)$,…,and $F(t\text{-}n)$, $F(t)$ is called a n-order fuzzy time series, and it can be represented by $F(t\text{-}n)$,…, $F(t\text{-}2)$, $F(t\text{-}1)\rightarrow F(t)$.

Definition 4:

If $F_1(t)$ is caused by ($F_1(t\text{-}1)$, $F_2(t\text{-}1)$), ($F_1(t\text{-}2)$, $F_2(t\text{-}2)$),…, ($F_1(t\text{-}n)$, $F_2(t\text{-}n)$), $F_1(t)$ is called a n-order 2-factor fuzzy time series, which is represented by ($F_1(t\text{-}n)$, $F_2(t\text{-}n)$),…, ($F_1(t\text{-}2)$, $F_2(t\text{-}2)$), ($F_1(t\text{-}1)$, $F_2(t\text{-}1)$)$\rightarrow F_1(t)$. Let $F_1(t)=X_t$ and $F_2(t)=Y_t$, where X_t and Y_t are fuzzy variables whose values are possible fuzzy sets of the first factor and the second factor, respectively, on day t. Then, a n-order 2-factor fuzzy logic relationship (FLR) (Chen, 2002) can be represented as follows,

$$(X_{t-n}, Y_{t-n}), \ldots, (X_{t-2}, Y_{t-2}), (X_{t-1}, Y_{t-1}) \rightarrow X_t,$$

where: (X_{t-n}, Y_{t-n}), …, (X_{t-2}, Y_{t-2}) and (X_{t-1}, Y_{t-1}), are referred to as the left-hand side (LHS) of the fuzzy logic relationship, and X_t is referred to as the right-hand side (RHS) of the fuzzy logic relationship.

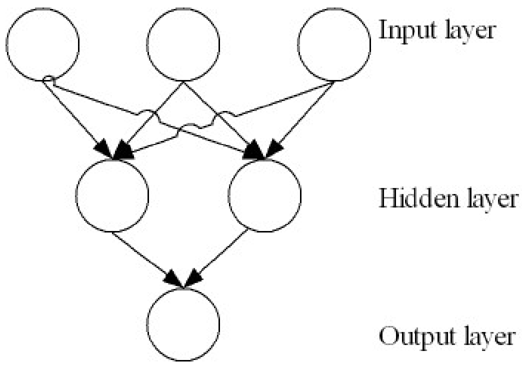
2.2 The Radial Basis Function Neural Networks

The concept of the ANN model was first introduced in 1950s (Leu *et al.*, 2009). Recently, many different ANN models have been proposed. Among them, the feed-forward neural network, back propagation neural network and the radial basis function neural network are the most well-known ANN models (Leu *et al.*, 2009; Panda and Narasimhan, 2007; Shih *et al.*, 2011). The framework of an ANN model is shown in Figure 1. It contains three layers: the input layer, the hidden layer and the output layer. This paper uses the radial basis function neural network model as the forecasting model because the radial basis function neural network model has been successfully applied in many financial applications (Panda and Narasimhan, 2007).

The input layer of a radial basis function neural network model broadcasts the coordinates of the input vector to each of the units in the hidden layer. Each unit in the hidden layer then produces an activation based on the associated radial basis function neural network model. Finally, each unit in the output layer computes a linear combination of the activations from the hidden units (Orr, 1996; Oyang *et al.*, 2005).

Figure 1

The Framework of an Artificial Neural Network



2.3 The Bootstrap Method

The bootstrap method was first proposed in 1979 (Efron, 1979). The bootstrap method is a statistical method to assign accuracy of measures to estimate samples (Diaconis and Efron, 1983). Generally, a bootstrap method is classified into the broader class of resampling methods. It can be implemented to generate a large number of resamples of the original dataset, each of which is obtained by random sampling with replacement from the original dataset. Subsequently, a particular statistic can be calculated from the collected values of the sampling distribution. Through simulations, it was found that the bootstrap method provides less biased statistics (Efrona and Tibshirani, 1986). Hence, the bootstrap method can be used to enhance the measures of statistical accuracy (Lajbcygier and Conner, 1997).

3. The Proposed Model

In this paper, the proposed model was modified from our previous work, namely the WFTSNN model (Leu *et al.*, 2011). There are three main differences between the proposed model and the WFTSNN model. (1) The WFTSNN model only gave different weights in two factors. However, due to the fact that the effect of the orders in the LHS of the FLR are different, we give different weights for the orders in the LHS of the FLR to enhance the performance in searching similar FLRs from the historical data. (2) The second factor of the WFTSNN used the spot price. However, the option price should be affected by the spot price and the strike price. Accordingly, we use the ratio of S/K (S is the spot price; K is the strike price) as the second factor in this paper. (3) Finally, because the option data is usually a small number of sample size dataset, the bootstrap method is used to enhance the forecasting accuracy.

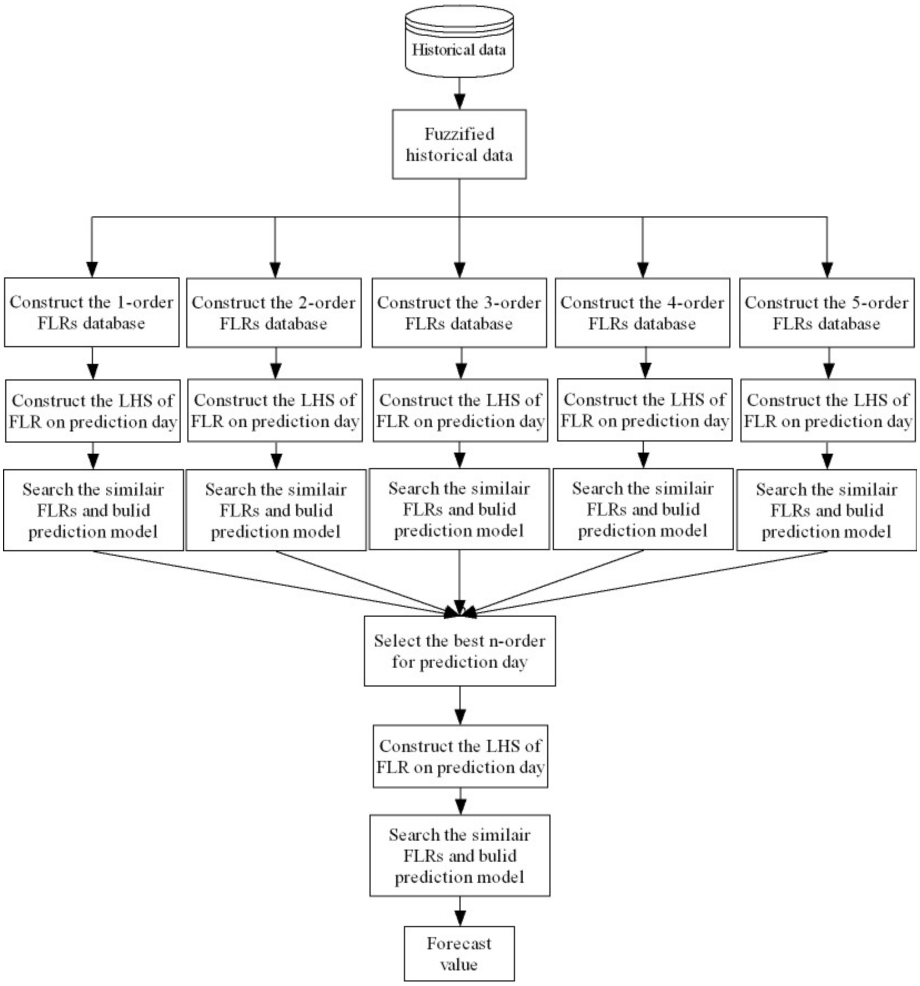
The proposed model includes three parts. Firstly, a dynamic n-order 2-factor fuzzy time series model can automatically choose the best n-order and use the selected n-order to search the similar FLRs from the historical data to generate a training FLRs database. Secondly, the radial basis function neural network model is used to build a prediction model by using the selected training dataset. Finally, the bootstrap method is used to enhance the forecasting accuracy. However, when there are many orders in the LHS of a FLR, it is difficult to find a matched FLR for prediction. To counter this problem, we refer to the literatures (Leu *et al.*, 2009; Yu, 2005) and use two ideas to search the similar FLRs in the second part of the proposed model. Thirdly, owning that the number of searched similar FLRs is too small, the bootstrap method is used to enhance the prediction accuracy for training with small number of samples. The flowchart of the proposed model is shown in Figure 2. The detailed procedure of the proposed model is described in the following steps.

Step 1: Divide the universe of discourse.

The universe of discourse (UoD) of the first factor is defined as $U=[D_{min}-D_1, D_{max}+D_2]$, where D_{min} and D_{max} are the minimum and maximum of the first factor, respectively; D_1 and D_2 are two positive real numbers to divide the UoD into n equal length intervals. The UoD of the second factor is defined as $V=[V_{min}-V_1, V_{max}+V_2]$, where V_{min} and V_{max} are the minimum and maximum of the second factor, respectively; similarly, V_1 and V_2 are two positive real numbers used to divide the UoD of the second factor into m equal length intervals. In this paper, we choose the call option closing price of TXO as the first factor and the ratio of S/K (S is the spot price; K is the strike price) as the second factor. We give an example to explain how to divide the universe of discourse. Suppose the maximum call option closing price of TXO and the minimum call option closing price of TXO were 1439 and 13, respectively, in 2005. If the length of intervals was set at 10, then D_{min}=13, D_1=3, D_{max}=1439 and D_2=1.

Figure 2

The Flowchart of the Proposed Model



<u>Step 2</u>: Define the fuzzy sets.

Linguistic terms A_i, $1 \le i \le n$, are defined as fuzzy sets on the intervals of the first factor. They are defined as follows:

$$A_1 = 1/u_1 + 0.5/u_2 + 0/u_3 + \cdots + 0/u_{n-2} + 0/u_{n-1} + 0/u_n,$$
$$A_2 = 0.5/u_1 + 1/u_2 + 0.5/u_3 + \cdots + 0/u_{n-2} + 0/u_{n-1} + 0/u_n,$$
$$\vdots$$
$$A_{n-1} = 0/u_1 + 0/u_2 + 0/u_3 + \cdots + 0.5/u_{n-2} + 1/u_{n-1} + 0.5/u_n,$$
$$A_n = 0/u_1 + 0/u_2 + 0/u_3 + \cdots + 0/u_{n-2} + 0.5/u_{n-1} + 1/u_n,$$

where: u_i denotes the ith interval of the first factor.

Similarly, linguistic term B_j, $1 \le j \le m$, is defined as a fuzzy set on the intervals of the second factor. They are defined as follows:

$$B_1 = 1/v_1 + 0.5/v_2 + 0/v_3 + \cdots + 0/v_{m-2} + 0/v_{m-1} + 0/v_m,$$
$$B_2 = 0.5/v_1 + 1/v_2 + 0.5/v_3 + \cdots + 0/v_{m-2} + 0/v_{m-1} + 0/v_m,$$
$$\vdots$$
$$B_{m-1} = 0/v_1 + 0/v_2 + 0/v_3 + \cdots + 0.5/v_{m-2} + 1/v_{m-1} + 0.5/v_m,$$
$$B_m = 0/v_1 + 0/v_2 + 0/v_3 + \cdots + 0/v_{m-2} + 0.5/v_{m-1} + 1/v_m,$$

where: v_i is the ith intervals of the second factor.

Step 3: *Determine the best order of the proposed model.*
The proposed model uses a dynamic *n*-order 2-factor fuzzy time series model to search the similar FLRs of the prediction day from the historical data to generate a training dataset for building the prediction model. The similar FLRs imply the similar trends of the prediction day in the historical data. To determine a suitable length of the trend will enhance the accuracy of the prediction model. In the proposed model, the number of orders can be regarded as the length of the trends. Hence, how to determine a suitable *n*-order is a challenge to the proposed model. To solve this problem, we build five different prediction models with different orders, which are set from 1 to 5 orders, respectively. Then, we choose the one with the best training accuracy to determine the best *n*-order. One should note that since *n* can be equal to any of 1, 2, 3, 4, or 5, Step 3(a) to Step 3(d) in the following will be performed five times to build five models with different orders.

(a) Construct the historical FLRs database with *n*-order.
For the historical data on day *i*, let X_{i-n}, Y_{i-n} denote the fuzzy set of $F_1(i-n)$ and $F_2(i-n)$ of the fuzzy time series. Let X_i denotes the fuzzy set of F_1(i). The FLRs with *n*-order on day *i* can be represented as follows:

$$(X_{i-n}, Y_{i-n}), \ldots, (X_{i-2}, Y_{i-2}), (X_{i-1}, Y_{i-1}) \rightarrow X_i.$$

Table 1 is an example of the historical FLRs database with *3*-order.

Table 1

An Example of the Historical FLRs Database with *3*-order

FLR	LHS		RHS
FLR_1	$(X_1, Y_2), (X_2, Y_2), (X_3, Y_3)$	→	X_4
FLR_2	$(X_2, Y_2), (X_3, Y_3), (X_4, Y_4)$	→	X_5
…	…	…	…
RLR_m	$(X_{m-3}, Y_{m-3}), (X_{m-2}, Y_{m-2}), (X_{m-1}, Y_{m-1})$	→	X_m

(b) Construct the LHS of FLR with *n*-order on the predicting day.
The LHS of the FLR with *n*-order on day *t* can be represented as follows:

$$(X_t\text{-}n, Y_t\text{-}n), \ldots, (X_t\text{-}2, Y_t\text{-}2), (X_t\text{-}1, Y_t\text{-}1).$$

The LHS of the FLR with n-order on day *t* is called the LHS of the prediction day in the following and is used to search the similar FLRs in Step 3(c).

(c) Search the similar FLRs to generate a training dataset.

Due to the fact that the effect of the orders in the LHS of the FLR is different, we give different weights to the orders in the LHS of the FLR. After setting the weights of each order, we calculate the weighted Euclidean distance (*WED*) of the LHS of the prediction day against the LHS of each FLR in the historical FLRs database. Subsequently, we select the top six similar FLRs with the smallest weighted Euclidean distance from the historical FLRs database as the training dataset to build the radial basis function neural network for forecasting. However, each of two factors of fuzzy time series plays a different role for prediction. In the proposed model, the first factor is more important than the second factor; we, therefore, assign a higher weight to the first factor to calculate the weighted Euclidean distance. The weighted Euclidean distance between the LHS of the prediction day and the LHS of the *i*th FLR in the historical FLRs database can be calculated according to equations (1)-(3).

$$WED_{Ai} = \sqrt{\sum_{j=1}^{n} j(PX_{t-j} - CX_{i-j})^2}, \tag{1}$$

$$WED_{Bi} = \sqrt{\sum_{j=1}^{n} j(PY_{t-j} - CY_{i-j})^2}, \tag{2}$$

$$WED_{i} = \frac{2 \times ED_{Ai} + ED_{Bi}}{3}. \tag{3}$$

where: PX_{t-j} and PY_{t-j} are the fuzzy sets' subscripts of the first factor and the second factor, respectively, of the *j*th order on the prediction day. Similarly, CX_{i-j} and CY_{i-j} are the fuzzy sets' subscripts of the first factor and the second factor, respectively, of the *j*th order of the *i*th FLR in the historical FLRs database.

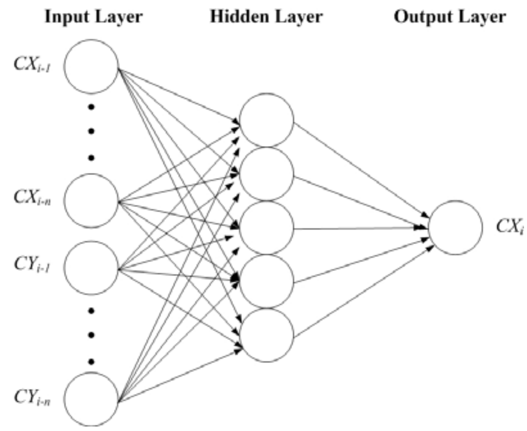
(d) Build prediction radial basis function neural network models with the bootstrap method.

With the top six similar FLRs training dataset, we can train a radial basis function neural network model for forecasting. However, the number of similar FLRs is too small. Hence, in this paper we use the bootstrap method to counter this problem and to enhance the prediction accuracy. In this step, we perform 100 times bootstrap methods. Namely, we perform 100 times radial basis function neural network model to generate 100 predictions. Subsequently, the average of 100 prediction results is as the final training prediction.

The training framework of the radial basis function neural network model is shown in Figure 3. The inputs are the LHS of the FLRs. The output is the RHS of the FLRs. Simply, the *1*st to the *n*th input variables are the subscripts of fuzzy sets of LHS's *1*st factor of the FLRs and the (*n+1*)th to the *2n*th input variables are the subscripts of fuzzy sets of LHS's *2*nd factor of the FLRs.

Figure 3

The Training Framework of the Neural Network Model



(e) Select the best *n*-order for prediction day by leave-one-out cross-validation.

Once the five models are built, the best model will be selected to determine the best *n*-order for prediction. In this paper, we use the prediction error of the model, which is calculated by leave-one-out cross-validation (LOOCV), as the criterion of model selection. Hence, the LHS of the FLR is used as the input of a trained model and the prediction error is calculated according to equation (4). The model with the smallest error will be selected for the prediction day. One should note that in equation (4) the FRHS denotes the subscript of the forecasted fuzzy set of FLR, and the TRHS denotes the actual subscript of the fuzzy set for the RHS of FLR.

$$error = \sum_{i=1}^{6} \left| FRHS_i - TRHS_i \right| \tag{4}$$

<u>Step 4</u>: Forecasting day *t*.

When the best *n*-order is determined in Step 3, we firstly re-build the training model with the best *n*-order by the selected similar FLRs, and perform forecasting by feeding the LHS of the FLR on the predicting day into the constructed radial basis function neural network with bootstrap model to get the forecasted subscript of the RHS on the prediction day. Because the forecasted value is a subscript of a fuzzy set, we have to defuzzify it into the option price forecasting value. We use the weighted average method as the defuzzification method. Equation (5) shows the weighted average defuzzification method.

$$forecast\_value = \begin{cases} \dfrac{M[1] + 0.5 \times M[2]}{1 + 0.5} & k = 1, \\ \dfrac{0.5 \times M[k-1] + M[k] + 0.5 \times M[k+1]}{0.5 + 1 + 0.5} & 1 < k \le n-1, \\ \dfrac{0.5 \times M[n-1] + M[n]}{0.5 + 1} & k = n, \end{cases} \tag{5}$$

where: *M*[*k*] denotes the midpoint value of the fuzzy set *k*. Note that an iteration of the above procedure (Step 1 through Step 4) predicts only one forecasting value.

4. Results and Performance

4.1 The Dataset

The dataset in this paper are the daily transaction data of TXO and TAIEX from January 3, 2005 to December 29, 2006. This paper investigates a sample of 23,819 call option price data. Call options can be divided into three moneyness categories according to their *S*/*K* ratio. The distribution of the dataset of three moneyness categories is shown in Table 2. In Table 2, we refer to the literature (Tseng *et al.*, 2008; Wang, 2009) for the definition of the moneyness categories. The dataset comprises 30 different strike price from 5,200 to 8,200 and 12 different expiration dates from January 2005 to December 2006. Note that the option prices of the beginning 10 transaction dates of each option are not predicted due to insufficient historical data. In predicting the option price of a specific date, the option prices of the previous transaction dates become the historical data.

Table 2

Data Distribution According to Categories of Moneyness

Categories	Moneyness	Number
In-the-money	$S/K > 1.02$	8938
At-the-money	$0.95 < S/K \leqq 1.02$	7508
Out-of-the-money	$S/K \leqq 0.95$	7373

Note: S is the spot price; K is the strike price.

4.2 Performance Measures

There are many measures of forecast accuracy (Hyndman and Koehler, 2006). To compare the performance of the proposed model with that of the existing models, we choose the same performance measures, which are used in the literature. Two different performance measures, the mean absolute error (MAE) and the root mean square error (RMSE), are used to measure the prediction accuracy of the proposed model and that of the existing models. The equations are shown in the following:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(A_t - P_t)^2}{n}}, \qquad 6)$$

$$MAE = \frac{\sum_{t=1}^{n}\left|A_t - P_t\right|}{n}, \qquad (7)$$

where: A_t and P_t denote the actual option price and the forecasting option price on day *t*, respectively.

4.3 Performance

The performance of the proposed model is compared with the other existing methods which were published in the previous literature (Leu *et al.*, 2010, 2011, Tseng *et al.*, 2008, Wang, 2009). Table 3 shows part of the results of the option price forecasting by the proposed model for an option with strike price equal to 6,000 and expiration date in April 2005. According to Table 3, the forecasting prices close to the actual option prices, except some dates when the option prices change abruptly. The performance of the proposed model and the other existing models are shown in Table 4. It shows that the performance of the proposed model is better than the other models besides the in-the-money category. Table 4 also shows that the performance of a traditional method, namely GARCH, is always the worst. The other hybrid models are better than GARCH, especially our proposed model. Furthermore, Figure 4 shows the forecasting results of an option with strike price equal to 7,400 and expiration date in December 2006. In Figure 4, the forecasting option price of the proposed model is closer to the actual option price.

Table 3

Part of the Results of the Proposed Model for an Option with Strike Price Equal to 6,000 and Expiration Date in April 2005

Transaction date	Actual option price	The proposed model forecasting
2005/02/15	232	235
2005/02/16	222	235
2005/02/17	180	225
2005/02/18	203	175
2005/02/21	226	205
2005/02/22	190	215
2005/02/23	207	215
2005/02/24	200	195

Table 4

The Performance in RMSE and MAE

	RMSE				MAE		
	Category				Category		
Methods	In	At	Out		In	At	Out
The proposed model	70.42	25.13*	5.30*		54.58	16.72*	3.24*
WFTSNN	67.94*	29.72	8.96		52.13	19.99	6.05
FTSNN	72.79	36.99	16.19		52.11*	23.98	9.78
GARCH	85.49	44.02	25.73		69.54	34.73	18.78
GJR	76.19	41.06	25.53		59.28	31.67	17.41
Gery-GJR	73.76	40.11	25.89		56.13	30.21	17.26
EGARCH	73.90	41.35	26.34		57.02	32.17	18.30
Gery-EGARCH	72.11	40.26	26.13		57.26	32.51	18.26

*Notes: 1. * denotes the smallest value. 2. GJR denotes GJR–GARCH model. 3. Grey-GJR denotes Grey-GJR–GARCH model.*

Although the performance of the proposed model in the in-the-money category is lower than that of WFTSNN and that of FTSNN, their performance is close. To compare the performance of the proposed model, the WFTSNN, and the FTSNN, we use the Diebold-Marioano test, which is a testing method to compare the forecast accuracy of two forecast models to evaluate their performance, as shown in Table 5. The values in Table 5 denote the *p*-value of the Diebold-Marioano test. Table 5 shows that the performance of WFTSNN and that of FTSNN is insignificantly different in all categories at α=0.05. In contrast, the performance of the proposed model is significantly different belonging to the in-the money category at α=0.05. Furthermore, in the other categories, the performance of the proposed model is significantly different at α=0.01. Hence, the proposed model offers a useful alternative for option price forecasting because it brings better performance in terms of the testing the results of the Diebold-Marioano test.

Figure 4

Time Series of the Actual Option Price and the Forecasting Price of an Option with Strike Price Equal to 7,400 and Expiration Date in December 2006

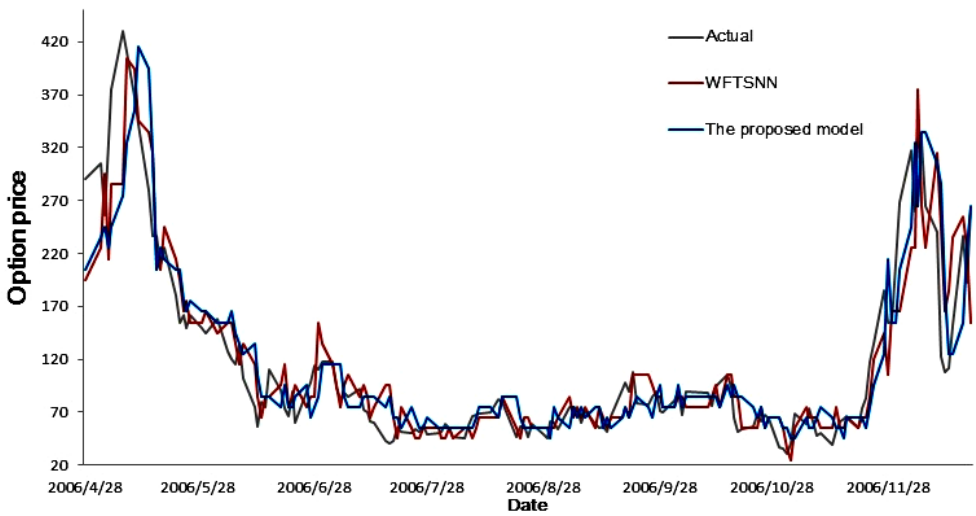


Table 5

The Diebold-Mariano Test for the Proposed Model, WFTSNN, and FTSNN

	Category				
	In			At	
	WFTSNN	FTSNN		WFTSNN	FTSNN
The proposed model	0.0468*	0.0413*		<0.001**	0.0015**
WFTSNN		0.7383			0.3393
	Category				
	Out			All	
	WFTSNN	FTSNN		WFTSNN	FTSNN
The proposed model	<0.001**	<0.001**		<0.001**	<0.001**
WFTSNN		0.553			0.8374

*Notes: 1. * denotes significance at α=0.05. 2. ** denotes significance at α=0.01*

5. Conclusion

Recently, the option price forecasting has become an important financial research issue. The option price forecasting is still a challenging problem due to the fact that it is affected by many factors. To counter this problem, in this paper we propose a novel hybrid model to predict the option price. In the proposed model, the dynamic *n*-order *2*-foctor fuzzy time series model can automatically choose the best *n*-order and use the selected *n*-order to search the similar FLRs to generate a training dataset from the historical FLRs database for training radial basis function neural network model for forecasting option price. Finally, the bootstrap method is used to enhance the prediction accuracy because the training dataset in this paper is a small number of samples.

The experiment results show that the proposed model is more accurate than the other existing methods in terms of RMSE and MAE for options belonging to at-the-money and out-the-money categories. Moreover, the proposed model brings the better forecasting accuracy after the Diebold-Marioano test. Hence, the proposed model offers a useful alternative for option price forecasting.

References

Alhassan, J. K., Misra, S., 2011. Using a weightless neural network to forecast stock prices: A case study of Nigerian stock exchange. *Scientific Research and Essays*, 6, pp. 2934-2940.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81, pp. 637.

Chen, S.-M., 2002. Forecasting enrollments based on high-order fuzzy time series. *Cybernetics and Systems*, 33, pp. 1-16.

Cheng, C.-H., Chen, T.-L., Teoh, H. J., Chiang, C.-H., 2008. Fuzzy time-series based on adaptive expectation model for TAIEX forecasting. *Expert Systems with Applications*, 34, pp. 1126-1132.

Diaconis, P., Efron, B., 1983. Computer-intensive methods in statistics. *Scientific American*, pp. 116-130.

Efron, B., 1979. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7, pp. 1-26.

Efron, B., Tibshirani, R., 1986. Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statistical Science*, 1, pp. 54-75.

Grudnitski, G., Osburn, L., 1993. Forecasting S&P and gold futures prices: An application of neural networks. *Journal of Futures Markets*, 13, pp. 631-643.

Huang, S.-C., 2008. Online option price forecasting by using unscented Kalman filters and support vector machines. *Expert Systems with Applications*, 34, pp. 2819-2825.

Hyndman, R.J., Koehler, A.B., 2006. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), pp. 679-688.

Ko, P.-C., 2009. Option valuation based on the neural regression model. *Expert Systems with Applications*, 36, pp. 464-471.

Lajbcygier, P., 2004. Improving option pricing with the product constrained hybrid neural network. *IEEE Transactions on Neural Networks*, 15, pp. 465-476.

Lajbcygier, P. R., Conner, J. T., 1997. Improved option pricing using bootstrap methods, *International Conference on Neural Networks*, 2194, pp. 2193-2197.

Lee, L.-W., Wang, L.-H., Chen, S.-M., Leu, Y., 2006. Handling forecasting problems based on two-factor high-order fuzzy time series. *IEEE Transactions on Fuzzy Systems*, 14, pp. 468-477.

Leu, Y., Lee, C.-P., Hung, C.-C., 2010. A fuzzy time series-based neural network approach to option price forecasting, *Proceedings of the Second international conference on intelligent information and database systems*, pp. 360-369.

Leu, Y., Lee, C.-P., Hung, C.-C., 2011. *A weighted fuzzy time series based neural network approach to option price forecasting*. Granular Computing and Intelligent Systems, Springer Berlin Heidelberg, pp. 237-248.

Leu, Y., Lee, C.-P., Jou, Y.-Z., 2009. A distance-based fuzzy time series model for exchange rates forecasting. *Expert Systems with Applications*, 36, pp. 8107-8114.

Mercuri, L., 2008. Option pricing in a Garch model with tempered stable innovations. *Finance Research Letters*, 5, pp. 172-182.

Morelli, M. J., Montagna, G., Nicrosini, O., Treccani, M., Farina, M., Amato, P., 2004. Pricing financial derivatives with neural networks. *Physica A*, 338, pp. 160-165.

Morariu, N., Iancu, E., Vlad, S., 2009. A Neural Network Model for Time-Series Forecasting. *Romanian Journal of Economic Forecasting*, 12(4), pp. 213-223.

Oyang, Y.-J., Hwang, S.-C., Ou, Y.-Y., Chen, C.-Y., Chen, Z.-W., 2005. Data classification with radial basis function networks based on a novel kernel density estimation algorithm. *IEEE Transactions on Neural Networks*, 16(1), pp. 225-236.

Orr, M.J.L., 1996. *Introduction to radial basis function networks*: Center Cognitive Sci., Univ. Edinburgh.

Panda, C., Narasimhan, V., 2007. Forecasting exchange rate better with artificial neural network. *Journal of Policy Modeling*, 29, pp. 227-236.

Saman, C., 2011. Scenarios of the Romanian GDP Evolution with Neural Models. *Romanian Journal for Economic Forecasting*, 14, pp. 129-140.

Shin, T., Han, I., 2000. Optimal signal multi-resolution by genetic algorithms to support artificial neural networks for exchange-rate forecasting. *Expert Systems with Applications*, 18, pp. 257-269.

Shih, K.-H., Cheng, C.-C., Wang, Y.-H., 2011. Financial Information Fraud Risk Warning for Manufacturing Industry - Using Logistic Regression and Neural Network. *Romanian Journal of Economic Forecasting*, 14(1), pp. 54-71.

Song, Q., Chissom, B. S., 1993. Forecasting enrollments with fuzzy time series - Part I. *Fuzzy Sets and Systems*, 54, pp. 1-9.

Song, Q., Chissom, B. S., 1993. Fuzzy time series and its models. *Fuzzy Sets and Systems*, 54, pp. 269-277.

Song, Q., Chissom, B. S., 1994. Forecasting enrollments with fuzzy time series - part II. *Fuzzy Sets and Systems*, 62, pp. 1-8.

Takahama, T., Sakai, S., Hara, A., Iwane, N., 2009. Predicting stock price using neural networks optimized by differential evolution with degeneration. *International Journal of Innovative Computing, Information and Control*, 5, pp. 5021-5031.

Teoh, H. J., Chen, T.-L., Cheng, C.-H., 2007. Frequency-weighted fuzzy time-series based on Fibonacci sequence for TAIEX forecasting, *Proceedings of the 2007 international conference on Emerging technologies in knowledge discovery and data mining*, pp. 27-34.

Tseng, C.-H., Cheng, S.-T., Wang, Y.-H., Peng, J.-T., 2008. Artificial neural network model of the hybrid EGARCH volatility of the Taiwan stock index option prices. *Physica A*, 387, pp. 3192-3200.

Wang, C.-C., Tu, Y.-H., Wong, H.-L., 2009. The comparison between multivariate fuzzy time series and traditional time series modeling to forecasting China exports, *Proceedings of the 2009 Fourth International Conference on Innovative Computing, Information and Control*, pp.1487-1490.

Wang, Y.-H., 2009. Nonlinear neural network forecasting model for stock index option price: Hybrid GJR-GARCH approach. *Expert Systems with Applications*, 36, pp. 564-570.

Yu, H.-K., 2005. Weighted fuzzy time series models for TAIEX forecasting. *Physica A*, 349, pp. 609-624.