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
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 model 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-factor 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 \( Y(t) = \ldots, 0, 1, 2, \ldots \), a subset of \( R \), be the universe of discourse in which fuzzy sets \( f_i(t) \) \((i = 1, 2, \ldots)\) 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 \( f_i(t) \in F(t) \), there is a \( f_i(t-1) \in F(t-1) \), such that there is a fuzzy relation \( R_{ij}(t, t-1) \) and \( f_i(t) = f_i(t-1) \), \( R_{ij}(t, t-1) \) where ‘\( \cdot \)’ is the max-min composition, \( F(t) \) is said to be caused by \( F(t-1) \) only, and it can be represented by \( F(t-1) \rightarrow F(t) \).

Definition 3:
If \( F(t) \) is caused by \( F(t-1), F(t-2), \ldots, and F(t-n) \), \( F(t) \) is called a \( n \)-order fuzzy time series, and it can be represented by \( F(t-n), \ldots, F(t-2), F(t-1) \rightarrow F(t) \).

Definition 4:
If \( F_i(t) \) is caused by \( (F_i(t-1), F_i(t-2)), (F_i(t-2), F_i(t-2)), \ldots, (F_i(t-n), F_i(t-n)) \), \( F_i(t) \) is called a \( n \)-order 2-factor fuzzy time series, which is represented by \( (F_i(t-n), F_i(t-n)), \ldots, (F_i(t-2), F_i(t-2)), (F_i(t-1), F_i(t-1)) \rightarrow F_i(t) \). Let \( F_i(t) = X_i \) and \( F_i(t) = Y_i \), where \( X_i \) and \( Y_i \) 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: $(X_{t-n}, Y_{t-n}), \ldots, (X_{t-2}, Y_{t-2})$ 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).

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 (Efron 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 \( U=[D_{\text{min}}-D_1, D_{\text{max}}+D_2] \), where \( D_{\text{min}} \) and \( D_{\text{max}} \) 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 \( V=[V_{\text{min}}-V_1, V_{\text{max}}+V_2] \), where \( V_{\text{min}} \) and \( V_{\text{max}} \) 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 \( D_{\text{min}}=13, D_1=3, D_{\text{max}}=1439 \) and \( D_2=1 \).
Step 2: Define the fuzzy sets.

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

$$A_i = \frac{1}{u_i} + 0.5/u_2 + 0/u_3 + \cdots + 0/u_{n-2} + 0/u_{n-1} + 0/u_n,$$

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

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

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

where: $u_i$ denotes the $i$th interval of the first factor.
Similarly, linguistic term $B_j, 1 \leq j \leq m$, is defined as a fuzzy set on the intervals of the second factor. They are defined as follows:

\[
\begin{align*}
B_1 &= \frac{1}{\nu_1} + 0.5/\nu_2 + 0/\nu_3 + \cdots + 0/\nu_{m-2} + 0/\nu_{m-1} + 0/\nu_m, \\
B_2 &= 0.5/\nu_1 + 1/\nu_2 + 0.5/\nu_3 + \cdots + 0/\nu_{m-2} + 0/\nu_{m-1} + 0/\nu_m, \\
&\vdots \\
B_{m-1} &= 0/\nu_1 + 0/\nu_2 + 0/\nu_3 + \cdots + 0.5/\nu_{m-2} + 1/\nu_{m-1} + 0.5/\nu_m, \\
B_m &= 0/\nu_1 + 0/\nu_2 + 0/\nu_3 + \cdots + 0/\nu_{m-2} + 0.5/\nu_{m-1} + 1/\nu_m,
\end{align*}
\]

where: $\nu_i$ is the $i$th interval 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>
<thead>
<tr>
<th>FLR</th>
<th>LHS</th>
<th>RHS</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLR1</td>
<td>$(X_2, Y_2), (X_3, Y_3), (X_4, Y_4)$</td>
<td>$\rightarrow X_5$</td>
</tr>
<tr>
<td>FLR2</td>
<td>$(X_3, Y_3), (X_4, Y_4)$</td>
<td>$\rightarrow X_5$</td>
</tr>
<tr>
<td></td>
<td>$\ldots$</td>
<td>$\ldots$</td>
</tr>
<tr>
<td>RLRm</td>
<td>$(X_{m-3}, Y_{m-3}), (X_{m-2}, Y_{m-2}), (X_{m-1}, Y_{m-1})$</td>
<td>$\rightarrow X_m$</td>
</tr>
</tbody>
</table>

(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:

\[(X_{t-n}, Y_{t-n}), \ldots, (X_{t-2}, Y_{t-2}), (X_{t-1}, Y_{t-1}).\]
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).

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

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

\[
WED = \frac{2 \times ED_{Ai} + ED_{Bi}}{3}.
\]

where: \( PX_{i,j} \) and \( PY_{i,j} \) 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 1st to the \( n \)th input variables are the subscripts of fuzzy sets of LHS’s 1st factor of the FLRs and the \((n+1)\)th to the 2nth input variables are the subscripts of fuzzy sets of LHS’s 2nd factor of the FLRs.
(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.

\[
\text{error} = \sum_{i=1}^{6} |\text{FRHS}_i - \text{TRHS}_i|
\]  

(4)

**Step 4: 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.

\[
\text{forecast\_value} = \begin{cases} 
\frac{M[1] + 0.5 \times M[2]}{1 + 0.5} & \text{if } k = 1, \\
\frac{0.5 \times M[k - 1] + M[k] + 0.5 \times M[k + 1]}{0.5 + 1 + 0.5} & 1 < k \leq n - 1, \\
\frac{0.5 \times M[n - 1] + M[n]}{0.5 + 1} & k = n,
\end{cases}
\]

(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

<table>
<thead>
<tr>
<th>Categories</th>
<th>Moneyness</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-the-money</td>
<td>( S/K &gt; 1.02 )</td>
<td>8938</td>
</tr>
<tr>
<td>At-the-money</td>
<td>( 0.95 &lt; S/K \leq 1.02 )</td>
<td>7508</td>
</tr>
<tr>
<td>Out-of-the-money</td>
<td>( S/K \leq 0.95 )</td>
<td>7373</td>
</tr>
</tbody>
</table>

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:

\[
RMSE = \sqrt{\frac{\sum_{t=1}^{n} (A_t - P_t)^2}{n}},
\]

\[
MAE = \frac{\sum_{t=1}^{n} |A_t - P_t|}{n},
\]

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

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

### Table 4

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

**Notes:**
1. * denotes the smallest value.
2. GJR denotes 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 $\alpha=0.05$. In contrast, the performance of the proposed model is significantly different belonging to the in-the-money category at $\alpha=0.05$. Furthermore, in the other categories, the performance of the proposed model is significantly different at $\alpha=0.01$. 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.

**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**

<table>
<thead>
<tr>
<th>Category</th>
<th>In</th>
<th>At</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WFTSNN</td>
<td>FTSNN</td>
</tr>
<tr>
<td>The proposed model</td>
<td>0.0468*</td>
<td>0.0413*</td>
</tr>
<tr>
<td>WFTSNN</td>
<td>0.7383</td>
<td>0.3393</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Category</th>
<th>In</th>
<th>At</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WFTSNN</td>
<td>FTSNN</td>
</tr>
<tr>
<td>The proposed model</td>
<td>&lt;0.001**</td>
<td>&lt;0.001**</td>
</tr>
<tr>
<td>WFTSNN</td>
<td>0.553</td>
<td>0.8374</td>
</tr>
</tbody>
</table>

Notes: 1. * denotes significance at $\alpha=0.05$. 2. ** denotes significance at $\alpha=0.01$. 

---

Romanian Journal of Economic Forecasting – XVII (2) 2014
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-factor 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


