Using neural network for forecasting TXO price under different volatility models

Ching-Ping Wang a,1, Shin-Hung Lin b,2, Hung-Hsi Huang c,* , Pei-Chen Wu d

a Graduate Institute of Finance, Economics, and Business Decision, National Kaohsiung University of Applied Sciences, No. 415, Jiaong Rd., Sanmin District, Kaohsiung City 80778, Taiwan
b Department of Finance, National Yunlin University of Science & Technology, No. 123, University Rd., Section 3, Douliou City 64002, Taiwan
c Department of Banking and Finance, National Chiayi University, No. 580, Simin Rd., Chiayi City 60054, Taiwan
d Graduate Institute of Finance, National Pingtung University of Science and Technology, No. 1, Hseuhfu Rd., Neiup, Pingtung 91201, Taiwan

1. Introduction

The Taiwan options market was established by the Taiwan Futures Exchange (TAIFEX) on December 24, 2001. TAIFEX options (TXO) were the only options traded in 2001. Besides, other options were subsequently introduced to date, including TEO, TFO, STO, MSO, XIO, GTO and TGO.3 Trading volume in TXO contracts totaled 1.57 million during 2002, and rapidly increased approximately 96.9 million during 2006. Although a total of eight different options are currently traded on the TAIFEX, TXO options still comprise over 90% of trading volume. Consequently, accurately evaluating TXO prices is important from both academic and practical perspectives.

Black and Scholes (1973) developed a European option pricing formula (known as the BS formula), and several theoretical option pricing models have subsequently been suggested (Cox, Ross, & Rubinstein, 1979; Heston & Nandi, 2000; Rubinstein, 1994). The BS formula presents option price as depending on five factors, including the current price and return volatility of the underlying asset, strike price, risk-free interest rate and time to maturity. Volatility is the only one of these factors that is not directly observable. Consequently, the literature introduces various models for estimating volatility.

Theoretical option pricing models generally assume that the underlying asset return follows a normal distribution and the market is frictionless. Since these assumptions deviate from reality, the above option pricing models have difficulty in efficiently evaluating option prices. Consequently, it is desirable to consider an informal mathematic formula for option pricing, such as the neural network method. Consequently, this study aims to use neural network for forecasting TXO prices under different volatility models. The neural network structure used in this study is the backpropagation neural network (BPNN), since this is the most popular structure. The input neurons in BPNN include the five factors of the BS formula. Since the volatility is not observable, this study provides five volatility models, including HV (historical volatility), IV (implied volatility), DVF (deterministic volatility function), GARCH and GM-GARCH (Grey Model-GARCH) models. Additionally, the futures index is substituted for the spot index, since the spot excludes cash dividends.

© 2011 Elsevier Ltd. All rights reserved.
Historical volatility models are relatively simple and widely used, and estimate volatility using the standard deviation of historical asset returns. The BS formula is used to extract the implied volatility from market option prices (Schmalensee & Trippi, 1978). Gemmill (1986) and Fleming (1998) verified that the implied volatility model outperforms the historical volatility model in forecasting option prices. To capture the nonconstant relation between volatility and strike prices, Dumas, Fleming, and Whaley (1998) developed a deterministic volatility function model that performed well in predicting option prices. Furthermore, the GARCH (generalized autoregressive conditional heteroskedasticity) model devised by Bollerslev (1986) has been verified to have good explanatory power for financial asset returns volatility. Akgiray (1989) demonstrated that the GARCH(1,1) model outperforms alternatives in predicting stock return volatilities. Successful studies have applied the GARCH model to forecast stock or index return volatilities (Claessen & Mittnik, 2002; Corredor & Santamaría, 2004; Lamoureux & Lastrapes, 1993). Recently, Chen, Hsin, & Wu (2010) addressed that the GARCH model outperforms alternative in forecasting exchange rate volatility. Additionally, Kayacan, Ulutas, and Kaynak (2010) employed the grey model to predict exchange rates.

Although the BS formula is historic and familiar both to academics and industry insiders interested in option pricing, it frequently has inferior out-of-sample performance than the neural network method (Amilon, 2003; Anders, Korn, & Schmitt, 1998). Several investigations have verified that the backpropagation method outperforms the traditional option pricing model in terms of forecasting performance (Lachtermacher & Fuller, 1995; Maliar- ius & Salchenberger, 1996; Yao, Li, & Tan, 2000). Furthermore, Wang (2009a, 2009b) and Lin and Yeh (2009) applied the backpropagation neural network to price TXO call options and compared forecasting performances among various volatility models and moneynesses using RMSE, MAE and MAPE as the performance criterion. However, this study introduces the best forecasting performance ratio (BFPR) as a new performance measure for option pricing.

The remainder of this paper is organized as follows. Section 2 introduces the methodology of backpropagation neural network and the forecasting performance criteria of RMSE, MAE and MAPE. Next, Section 3 describes the volatility models, including HV, IV, DVF, GARCH and GM-GARCH models. Section 4 then demonstrates the empirical results, including data description, the best forecasting performance ratios for various classifications, and the interaction of volatility models and moneynesses. Finally, Section 5 presents conclusions.

2. Backpropagation neural network

The backpropagation neural network (simply BPNN) comprises an input layer, one or more hidden layers, and an output layer. Following Haykin (1999), this study employs BPNN with one hidden layer for pricing options. Since according to the BS formula option price is determined by the five factors, including current price of the underlying asset $S_0$, strike price $K, time$ to maturity $T$, risk-free rate $r$, and volatility $\sigma$, several studies use these five factors as the neurons in the input layer; for example, Lin and Yeh (2009) and Wang (2009a, 2009b). However, since the spot index in Taiwan stock market does not reflect cash dividends, the futures index $F$ is being substituted for $S_0$ in this study.

Fig. 1 illustrates the neural network structure with one hidden layer. The input layer contains the six neurons of $F, K, T, r, \sigma$, and $I_1, I_2, I_3$, respectively. The activation function connecting the input layer and hidden layers is the logistic function or hyperbolic tangent function. Accordingly, the $k$th neuron value in the hidden layer is

$$V_k = \begin{cases} \frac{1}{1 + e^{-x}} & \text{for the logistic function} \\ \frac{1}{2} \left( 1 + \frac{x}{\sqrt{\pi}} \right) & \text{for the hyperbolic function} \end{cases}$$

The constant 1 represents the threshold value. Since the volatility $\sigma$ cannot be observed, Section 3 of this study introduces five models for volatility estimation. Section 4 of this study compares the prediction performances of the adopted 2, 3 or 4 neurons.

The logistic function and hyperbolic function are quoted from Chapter 4 of Haykin (1999).
where \( w_{ik} \) denotes the weight connecting the \( i \)th neuron in the input layer and the \( k \)th neuron in hidden layer, and \( b_k \) is a threshold. The activation function connecting the hidden and output layers is linear, such that the estimated or predicted option price is calculated by

\[
\hat{c} = \sum_k w_{ik} v_k - b
\]

where \( W_k \) denotes the weight connecting to the \( k \)th neuron in the hidden layer, and \( b \) is a threshold. The estimation or prediction performance for the BPN is measured by the following error function:

\[
E = \frac{1}{2} \sum_j (c_j - \hat{c}_j)^2
\]

where \( c_j \) and \( \hat{c}_j \) represent the actual and estimated values for the \( j \)th observed option price. The Levenberg–Marquardt algorithm is used to minimize the error function and find these values of \( \hat{c}_j \).

Using the backpropagation neural network, this study respectively forecasts the call and put option prices. Market data are partitioned into two parts, training and testing sets. Similar to Wang (2009a, 2009b) and Lin and Yeh (2009), the training part contains the earlier 70% of the data and the testing part contains the remaining 30%. This study respectively adopts RMSE, MAE and MAPE, defined below to measure pricing errors, following to An-.

\[
\begin{align*}
\text{RMSE} & = \sqrt{\frac{1}{n} \sum_{j=1}^{n} (c_j - \hat{c}_j)^2} \\
\text{MAE} & = \frac{1}{n} \sum_{j=1}^{n} |c_j - \hat{c}_j| \\
\text{MAPE} & = \frac{1}{n} \sum_{j=1}^{n} \frac{|c_j - \hat{c}_j|}{c_j}
\end{align*}
\]

where \( n \) is the number of observations.

### 3. Volatility models

#### 3.1. Historical volatility and implied volatility

Let \( S_t \) denote the stock index at date \( t \), where the daily return \( R_t = S_t/S_{t-1} - 1 \). Referring to Anders et al. (1998), Yao et al. (2000) and Amilon (2003), the daily historical volatility is

\[
\sigma_t = \sqrt{\frac{1}{29} \sum_{d=1}^{30} (R_{t,d} - \bar{R})^2}, \quad \bar{R} = \frac{1}{30} \sum_{d=1}^{30} R_{t,d}
\]

The average number of annual trading days is 250 during 2008 and 2009. Accordingly, the annual historical volatility

\[
\sigma_{dt} = \sigma_t \sqrt{250}
\]

The spot index \( S_0 \) is substituted by the futures price \( F \), and the BS formula for the call option price \( c \) and the put option price \( p \) is as follows:

\[
\begin{align*}
c &= e^{-rT} \text{N}(d_1) - KN(d_2) \\
p &= e^{-rT} \text{N}(-d_2) - FN(-d_1) \\
d_1 &= \frac{\ln(F/K) + rT}{\sigma \sqrt{T}}, \quad d_2 = d_1 - \sigma \sqrt{T}
\end{align*}
\]

where \( N(\cdot) \) denotes the cumulative distribution function of the standard normal variable. Besides \( \sigma \), the other parameters in Eq. (7) can be observed from the option market. Consequently, the implied volatility can be obtained from Eq. (7).

#### 3.2. The deterministic volatility function

Rubinstein (1994) illustrated that the BS implied volatility may vary with different moneynesses. Moreover, Dumas et al. (1998) examined the prediction and hedging performances on a deterministic volatility function. Furthermore, Peña, Rubio, and Serna (1999) suggested that a nonconstant relation exists between the strike price and the implied volatility. These phenomena are termed the volatility “smile” or “smirk” (Zhang & Xiang, 2008). Accordingly, this study assumes that the volatility satisfies a deterministic function

\[
\sigma(K,F) = \sigma_0 + \alpha_1 (K/F) + \alpha_2 (K/F)^2
\]

where \( K \) represents the strike price. For the traded options at date \( t - 1 \), \( \sigma(K,F) \) represents the BS implied volatility. Subsequently, using all the traded options at date \( t - 1 \) as the sample, parameters \( \sigma_0, \alpha_1 \) and \( \alpha_2 \) are estimated by the OLS regression coefficients in Eq. (8). Next, the estimated volatility at date \( t \) is

\[
\hat{\sigma}_t = \hat{\sigma}_{o1} + \hat{\sigma}_{11} (K/F) + \hat{\sigma}_{21} (K/F)^2
\]

where \( \hat{\sigma}_{o1}, \hat{\sigma}_{11} \) and \( \hat{\sigma}_{21} \) denote the estimates of \( \sigma_0, \alpha_1 \) and \( \alpha_2 \), respectively.

#### 3.3. The GARCH(1,1) model

The GARCH(1,1) model is frequently applied to forecast stock index return volatility; for instance, Akgiray (1989) and Claessen and Mittnik (2002). Additionally, RiskMetrics by Morgan (1995) adopts GARCH(1,1) for estimating stock return volatility. Mathematically, the GARCH(1,1) model can be presented as follows:

\[
\begin{align*}
\sigma_t^2 &= \alpha_0 + \beta_1 \sigma_{t-1}^2 + \gamma \varepsilon_{t-1}^2 \\
R_t &= \bar{R} + \varepsilon_t
\end{align*}
\]

where the error term \( \varepsilon_t \sim N(0, \sigma_t^2) \), \( R_t \) and \( \bar{R} \) are daily and average returns, respectively, and \( \alpha, \beta, \gamma \) are constant parameters.

#### 3.4. GM-GARCH model

Deng (1982) first suggested using the grey system theory to manage a stochastic control system problem. Subsequently, grey system theory has been widely applied in both academic research and practical contexts. (Deng, 1989; Liu & Forrest, 2007; Wang & Liu, 2009) The grey model (GM) transfers disordered data to a regular series without the need to conduct numerous observations. First-order grey model with one variable (GM(1,1)) is the basic type of grey system. Recently, GM(1,1) has been applied to improve stock index forecasting performance (Chang & Tsai, 2008; Chen, Hardle, & Jeong, 2010) and option pricing (Lin & Yeh, 2009; Wang, 2009a, 2009b). Since GARCH(1,1) has good power to explain

---

8 The algorithm, developed by Levenberg (1944) and Marquardt (1963), efficiently solves the least-squares estimations of nonlinear parameters.

9 RMSE is a common measurement of pricing error used in the literature; for example, by Amilon (2003), Christoffersen and Jacobs (2004), and Gemmill (1986).

10 Trading data are collected from the beginning of 2008 to the end of 2009, and the samples for each year comprise 249 and 251 trading days, respectively.
