A Pragmatical Option Pricing Method Combining Black-Scholes Formula, Time Series Analysis and Artificial Neural Network

Kai Liu
Department of Finance, School of Economics
Siming S.Rd. 422 Xiamen University
Xiamen, China
Email: liu.kai@stu.xmu.edu.cn

Xiao Wang
Department of Finance, School of Economics
Siming S.Rd. 422 Xiamen University
Xiamen, China
Email: deshaw_wang@163.com

***Abstract*—Although many theoretical methods were developed to price various derivatives, pricing deviation still remains very high. This paper provides a pragmatical option pricing method by combining skewness and kurtosis adjusted Black-Scholes model of Corrado and Su, time series analysis and Artificial Neural Network (ANN). The empirical tests in FTSE 100 Index options show that pricing deviation calculated by adjusted Black-Scholes model is still high. After the model is modified by time series analysis and ANN methods, the pricing deviation is reduced, which is much smaller than the previous models. It is suggested that time series analysis and Artificial Neural Network methods can be used in the pragmatical work to make the pricing more fast and precise.**

***Keywords*-Adjusted Black-Scholes Model; Times Series Analysis; Artificial Neural Network; Option Pricing**

I. Introduction

The Black-Scholes (1973)[1] option pricing model is used to value a variety of derivatives and securities. Despite its prominent usefulness, some empirical researches found the formula had some deficiencies which may lead to inaccuracies in assets valuation.These inaccuracies are originated from the simplified set of assumptions, such as geometric Brownian motion of stock return, constant variance,without taxes or continuous trading on the underlying assets, etc, which are obviously violated in financial market.

Some subsequent researches relaxed the assumptions to make the formula more approximate to the real market. Merton (1976)[2] employed jump diffusion model to describe the movement of the underlying assets return; Hull and White (1987)[3] combined stochastic volatility to modify the assumption of constant variance and some later approaches have tended to combine fat-tailed independent shocks and time varying variance. Corrado and Su (1996)[4] extended the Black-Scholes model by introducing the non-normal skewness and kurtosis into the formula and the estimation indicated significant skewness and kurtosis implied by option prices. However, the results provided by these generalizations and extensions did not manage to be truly consistent with the market data. Moreover, many extensions are often too complicated to implement and have poor market performances.

Artificial Neural Network (ANN) is a promising way in modifying the option pricing model. The former extensions of option pricing models are strongly correlated with financial theories, however, some abnormal price behaviors still can not be explained by the existed financial theory so that the pricing deviations are still high. The ANN technique is not set out from financial theory, but can make an estimation with a black-box method via the input variables. What's more, the market is changing frequently because the attitudes toward the option price changes time to time (eg. Rubinstein, 1985 [5]) and many theoretical methods are stationary models which may fail to value the options precisely. With ANN technique, it can be trained frequently to adapt the rapidly changing market conditions.

Beside Neural Network technique, statistical method, like time series analysis, is combined to improve the model performance. By examining the correlation of pricing deviations in different periods, the result indicated that the correlation between current pricing deviation and first-order lag deviation is very strong, close to 0.9. With the combination of these two methods, pricing deviations can be reduced significantly and the time consumption is relatively small so that the model is more meaningful and pragmatic to the investors.

The remainder of the article is organized as follows.The methodology of this paper is presented in the section 2. The adjusted Black-Scholes model and implied information mining method is described in Section 3. The section 4 showed the Time Series analysis with FTSE 100 Index Options data. In the section 5, Artificial Neural Network technique is described and explained. The empirical result is reported in section 6. Finally, conclusions are offered in section 7.

II. Methodology

The option price, in theoretical, is the discounted expected return of the option claim. However, there are lots of uncertainty that impact the expectation, many previous researches assumed the uncertainty follows some rules and calculated the option price basing on the assumption. They assume an

978-1-4799-2548-3/13 $31.00
DOI 10.1109/CIS.2013.38

ideal world but there are no ideal world, and that is why option pricing are not always precise.

This paper trisects option pricing into C_{BS},C_{AR} and C_{ANN} and analyzes each part in order to make an accurate pricing. The trisection can be formulized as below :

$$C = C_{BS} + C_{AR} + C_{ANN} + \varepsilon \quad (1)$$

C_{BS} is the pricing in an ideal world, so $C - C_{BS}$ is the pricing deviation from ideal world to real world. C_{AR} analyze the deviation and find out a time series trend of such deviation, therefore $C - C_{BS} - C_{AR}$ represents the deviation which can not be explained by time series trends. Finally, C_{ANN} is a fitting function trying to map the market information to the deviation which can not be explained by time series trends. The details of C_{BS},C_{AR} and C_{ANN} will be discussed in the following sections.

III. The Adjusted Black-Scholes Model and Implied Information Mining

In this section, a traditional option pricing formula is introduced, which can get an analytical solution under a hypothetically ideal market, and this can be used in a data mining process that estimates the implied information including risk-free interest rate, market volatility, skewness and kurtosis.

The Black-Scholes model is first published by Fischer Black and Myron Scholes (1973)[1]. Their model is based on the assumption of a normal-distribution return of underlying option , however, empirical study (Vahamaa ,2003[6]) shows that the return has skewness and kurtosis default normality. Corrado and Su adjusted this model by adding terms for non-normal skewness and kurtosis to make the model more cohesive to the real world[4].

The formula deduced from the Corrado and Su's model can be represented as follow (European style call option):

$$C = SN(d_1) - Ke^{-rT}N(d_2) + \mu_3 Q_3 + (\mu_4 - 3)Q_4 \quad (2)$$

where:
$Q_3 = \frac{1}{6}S\sigma\sqrt{T}[(2\sigma\sqrt{T} - d)n(d_1) + \sigma^2(d_1)]$
$Q_4 = \frac{1}{24}S\sigma\sqrt{T}[(d_1^2 - 1 - 3\sigma\sqrt{T}d_2)n(d_1) + \sigma^3 T^{3/2}N(d_1)]$
$d_1 = \frac{ln(S/K)+(r+\sigma^2/2)T}{\sigma\sqrt{T}}, d_2 = d_1 - \sigma\sqrt{T}$

K represents the strike price, S is the price of underlying asset, T is the time to maturity date, σ is the volatility and r is the risk-free interest rate. In addition to the original BS formula, terms $\mu_3 Q_3$ and $(\mu_4 - 3)Q_4$ are added to measure the effects of non-normal skewness and kurtosis, which μ_3 represents skewness and μ_4 kurtosis.

This formula is the most basic consistent with the pricing model in equation (1), for it represents the intrinsic option value in an ideal world. Thus, let C_{BS} equal C in equation (2). The residual term $C_{AR} + C_{ANN} + \varepsilon$ in equation (1) can be understood as the pricing deviation between ideal world and real world and will be analyzed in the following section.

The adjusted BS formula is useful to capture implied information (risk-free interest rate, market volatility, skewness and kurtosis). Many previous researches of pricing utilized BS formula, and the parameter estimation is still a controversial issue. Some researches employed a constant parameter, say, volatility, or estimated the parameters in a period, say, 60 days. But the determination of the interval of period remains divergence. Hull and White regarded the parameters as a stochastic process and refined the Black-Scholes formula with the method (Hull, 1987 [3]). However, this begets more estimation work. In this paper, a simple and smart method is developed to estimate the parameters, and it has been proved to be well-performed(Mayhew, 1995 [7]). The implied parameters suggest the expected by market average parameters of the period till maturity.

Implied parameters are estimated by minimizing the sum of squared deviations between the observed market prices and the price calculated by BS formula in equation (2).

$$sq(r, \sigma, \mu_3, \mu_4) = \sum (C - C_{BS}(r, \sigma, \mu_3, \mu_4))^2 \quad (3)$$

To get the implied parameters, Lagarias's Nelder-Mead simplex algorithm (Lagarias, 1998[8]) is used to minimize $sq(r, \sigma, \mu_3, \mu_4)$ in equation (3) and hence get the parameters estimated as implied ones which makes sq minimized.

In this study, a generally accepted assumption is made that the parameters r, σ, μ_3, μ_4 are continuous when time interval tends to zero, which means that no jump behavior is observed. That is:

$$\lim_{t \to t_0} \sigma_t = \sigma_{t_0} \qquad \text{and so do } r, \mu_3, \mu_4$$

Thus, current value of r, σ, μ_3, μ_4 can be used in the pricing model, since values of these parameters in the last second can be always known and they are very close to the current ones. Under this consideration, current values (or a approximate one) of r, σ, μ_3, μ_4 are accessible if the time interval is set small enough.

IV. Time Series Analysis

After pricing options with the BS formula (2), the residual term of the BS pricing model, $C_{AR} + C_{ANN} + \varepsilon$, represents the pricing deviation between ideal world and real world. Conjecture is made, without difficulty, that the pricing deviation of a specific contract will not change much during a small period or follows a certain rule. It can be easily found that, things like transaction fee and margin bring similar deviation to the option price, which make the deviations strongly correlate. Besides the transaction fee and margin, there may be lots of similar factors that affect the option price, these will be reflected on the time series trends. Therefore, time series analysis is incorporated to catch the rule of pricing deviation.

To prove the conjecture, Figure 1 shows a strong relationship between the pricing deviation of two neighboring time ticks. (FTSE 100 index option daily close price from 2008/4/1 to 2008/6/30, including average 314 contracts per day). The correlation is 0.9298, which suggests that time series analysis can help a lot in this pricing model.

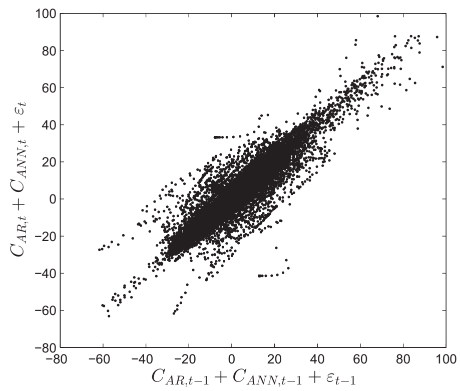

Figure 1. the correlation of pricing deviation between lags

Autoregressive-moving-average (ARMA) model is a statistic tool to analyze the time series data. It consists of two parts, the autoregressive(AR) part shows the relationship between the current data and the past data, the moving-average(MA) part shows the similar relationship between white noises. A general ARMA model (Whittle, 1951[9]) can be written as equation (4):

$$X_t = c + \epsilon_t + \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} \qquad (4)$$

ARMA model is applied in the pricing model described as equation 1 and let $C_{AR} + C_{ANN} + \varepsilon$ be X_t showed in equation (4). In practice, X_t is assigned as $C - C_{BS}$, every contract is considered respectively. ARMA model can analyze a potential statistical rule of the pricing deviation and a prediction can be made under the rule.

After the coefficients of ARMA model are estimated, X_t in equation (4) can be calculated through the estimated coefficients, X_{t-i} and ϵ_{t-i} used in the equation (4). The residual term of this equation, ϵ_t, can be understood as the pricing deviation which can be explained by time series trend. In equation (1), C_{BS} is the intrinsic option price assumed in an ideal world, and C_{AR} is the pricing deviation which can be explained by time series trend[1], so the residual term ϵ_t in equation (4) equals $C_{ANN} + \varepsilon$ in equation (1).

V. Artificial Neural Network

The previous 2 sections, *The Adjusted Black-Scholes Model and Implied Information Mining* and *Time Series Analysis*, discussed: (1) option pricing with theoretical formula in the ideal world; (2) mining the implied information hidden in the option price; (3) the time series trend of pricing deviation between ideal world and the real world. However, the pricing model is still inaccurate when these process completed.

So, what other factors determine the pricing deviation besides time series trends? There are lots of factors that impact the option price. However, such impacts may not have an intuitive formula. For example, the skewness is considered to be associated with market sentiment (Majmin, 2012 [10]), and this will surely affect the option price to some extent. Unfortunately, the sentiment or emotion effect is unmeasurable.

Now, let $C_{ANN}(information)$ be a function mapping $R^p \to R$.

$$C_{ANN} : information \to (C - C_{BS} - C_{AR}) \qquad (5)$$

This abstract function attempts to find the relationship between market information and the deviation which can not be explained by time series trend. Here, artificial neural network (ANN) is introduced to solve this fitting problem.Similar fitting problem (fitting residual of BS model) has been successfully solve by ANN before (Andreou, 2008 [11]).

The first thing to be considered is what the term $information$ consist of, in other word, what should be the inputs of function C_{ANN}. In the BS formula (2), the variables in the formula plays an important role. So these variables are important enough to included in $information$, another variable Δ [2] is included which is decisive in arbitrage process.

The variables included in the vector $information$ are listed below:

- S/K : the ratio of S and K
- T : time to maturity
- r : risk-free rate
- σ : volatility
- μ_3 : skewness
- μ_4 : kurtosis
- Δ : the change rate of an option price when the price of the underlying benchmark changes

Therefore, the $information$ is a vector of R^7, C_{ANN} is a function mapping $R^7 \to R$. A specific ANN, Cascade-forward network is used to fitting this function. The structure of Cascade-forward network is shown in Figure 2.

Finally, Levenberg-Marquardt (Hagan, 1999 [12]) algorithm can efficiently make the training of this networks convergence with second-order training speed.

[1] C_{AR}, instead of C_{ARMA}, is denominated because ARMA(1,0) is detected in the empirical study, which suggests an AR(1) model

[2] $\Delta = \frac{\partial C_{BS}}{\partial S} = N(d_1) + \mu_3 q_3 + (\mu_4 - 3)Q_4$ (Vahamaa, 2003 [6])
where
$q_3 = \frac{1}{6}[\sigma^3 T^{3/2} N(d_1) + (\frac{\phi_1 d_1}{\sigma\sqrt{T}} + \sigma^2 T - 1 - \phi_1) n(d_1)]$
$q_4 = \frac{1}{24}[\sigma^4 T^2 N(d_1) + \sigma^3 T^{3/2} n(d_1) + \frac{n(d_1)}{\sigma\sqrt{T}}(\phi_2 - 2\sigma^2 T + 2rT + 2ln(S/K)) - \frac{\phi_2 d_1 n(d_1)}{\sigma^2 T}]$
$\phi_1 = rT - (3/2)\sigma^2 T + ln(S/K)$
$\phi_2 = r^2T^2 - 2r\sigma^2T^2 + (7/4)\sigma^4T^2 - \sigma^2T + ln(S/K)(2rT - 2\sigma^2T + ln(S/K))$

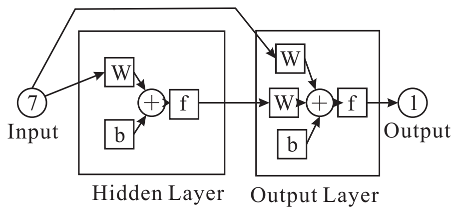

Figure 2. The Structure of Cascade-forward Network

VI. Empirical Study

A. Data

The data in this article contains the daily option close prices [3] of FTSE 100 index options traded at the London International Financial Futures and Options Exchange, aggregating 19833 samples. The sample period extends from April 4, 2008 to June 4, 2008, including average 314 contracts per day with different maturity days and strike prices.

B. Black-Scholes Pricing and Implied Information Mining

In the empirical study, the pricing and mining are processed day by day in order to figure out the daily implied information of the underlying FTSE 100 index. The implied information of FTSE 100 index from April 4, 2008 to June 4, 2008 is shown in Figure 3.

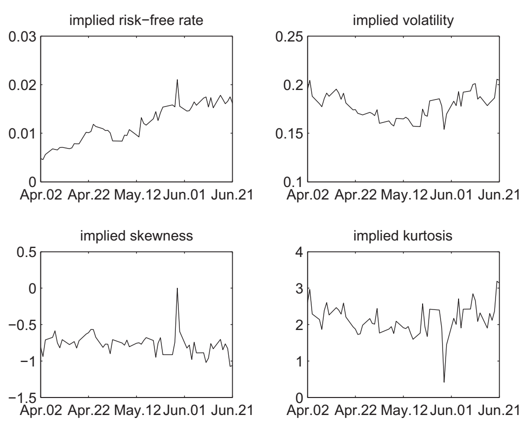

Figure 3. Implied Information

After the BS model is applied, the pricing performance and residuals are shown in Figure 4. Note that only a 1000 samples (5001 - 6000) is put in the figure in order to make it concise.

The average pricing deviation, which is measured in mean of absolute values, is 13.3618, and most of deviations spread from -50 to 50. So far the term C_{BS} in equation (2) is obtained. The pricing deviations in this subsection form

[3] minute price is also used in the empirical study and the pricing is almost exactly correct. Minute price tends to change a little which may not truely show how the model performs well, so daily data is put here to show the advantage of the model.

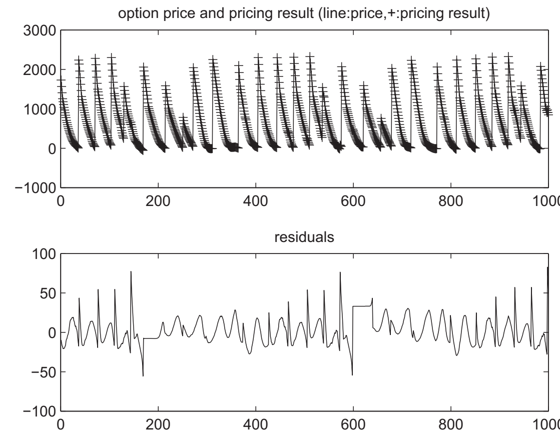

Figure 4. Pricing Performance

$C_{AR} + C_{ANN} + \varepsilon$, and will be analyzed in the following subsection. The average running time of this process is 0.23 second (Intel i5-3317U CPU, Windows 8 pro, Matlab R2013a).

C. Time Series Analysis

As mentioned in Section IV, the pricing deviation has some sorts of time series rule. The ARMA model is used here to detect the rule. Here, every time series data of single contract is considered respectively. Autocorrelation function (ACF) and Partial autocorrelation function (PACF) suggest that it should be ARMA(1,0) or simplified, AR(1).

In practice, each prediction is made using the time series data $[0, t-1]$, where t is the date under consideration. The result is shown in Figure 5. For the same reason, only 1000 samples (5001 - 6000) are put in the figure under the consideration of conciseness.

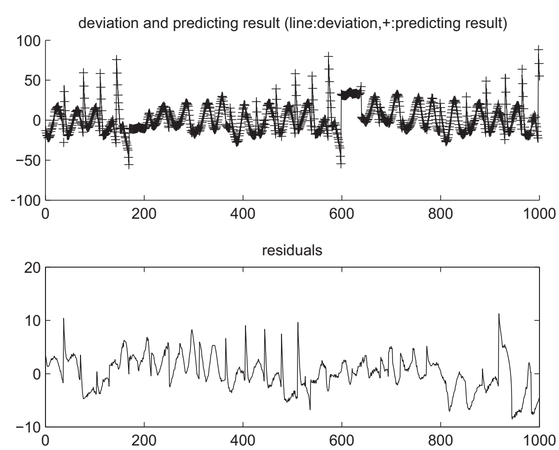

Figure 5. Predicting Performance

The average pricing residual is 2.5659 (mean of absolute values), and most of them spread from -10 to 10. C_{AR} has been calculated in this subsection, the last term, $C_{ANN} + \varepsilon$ will be analyzed in the next subsection. The average running time of this process is 0.00 second.

D. Artificial Neural Network

The remaining term $C_{ANN} + \varepsilon$ or equally, $C - C_{BS} - C_{AR}$ is difficult to be analyzed by theoretical explanation. However, it can be measured in a black box method, treated as a function which can be fitted. Artificial Neural Network, more specifically Cascade-forward network, is put here to fit the deviation which can not be explained by time series trends. [11] did a similar research using different pricing details and proved ANN is feasible on this matter, without incorporating time series analysis.

In practice, samples 1 - 5000 are used for learning and the neural network outputs fitting samples 5001 - 6000 basing on the learning. The performance is shown in Figure 6.

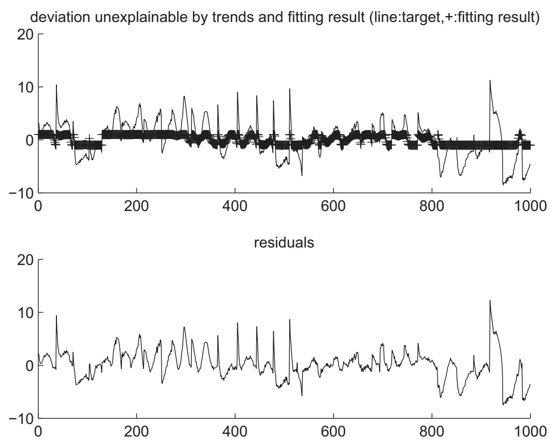

Figure 6. Fitting Performance

The picture shows that ANN can fit part of the deviations. Though most of them still spread from -10 to 10, the average pricing residual is reduced to 1.9766, which means that the pricing is more precise. The average running time of learning process is 10.23 seconds, however once the leraning process is completed, the average running time of fitting is 0.00 second, which is feasible in pratice.

VII. Conclusion

This paper aims at developing a pragmatical pricing method combining skewness and kurtosis adjusted Black-Scholes model of Corrado and Su (1996), time series analysis and Artificial Neural Network. The option price is trisected into BS theoretical part, AR time series trends part and ANN fitting part and each part is calculated one by one using the residual of the previous calculation. In the empirical study, the model analyzes the daily close price of FTSE 100 Index options, we can draw some conclusions. First, although the adjusted Black-Scholes has reduced pricing deviation to some extend, the deviation still remains very high. Second, modified the model with time series trend and ANN methods can make the option pricing become more accurate, without consuming too much time, during the calculations. It is suggested that our model is a more precise and timesaving method in pragmatical work. Futhermore, it can be found that the terms in equation (1), C_{BS}, C_{AR} and C_{ANN}, are first order differentiable to the underlying price, which implies that the model can be used in a delta arbitrage process.

Acknowledgements

This work is supported by the school of economics, Xiamen University.

References

[1] F. Black and M. Scholes, "The pricing of options and corporate liabilities," *The journal of political economy*, pp. 637–654, 1973.

[2] R. C. Merton, "Option pricing when underlying stock returns are discontinuous," *Journal of financial economics*, vol. 3, no. 1, pp. 125–144, 1976.

[3] J. Hull and A. White, "The pricing of options on assets with stochastic volatilities," *The journal of finance*, vol. 42, no. 2, pp. 281–300, 1987.

[4] C. J. Corrado and T. Su, "Skewness and kurtosis in s&p 500 index returns implied by option prices," *Journal of Financial research*, vol. 19, pp. 175–192, 1996.

[5] M. Rubinstein, "Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active cboe option classes from august 23, 1976 through august 31, 1978," *The Journal of Finance*, vol. 40, no. 2, pp. 455–480, 1985.

[6] S. Vähämaa, "Skewness and kurtosis adjusted black–scholes model: A note on hedging performance," *Finance Letters*, vol. 1, no. 5, pp. 6–12, 2003.

[7] S. Mayhew, "Implied volatility," *Financial Analysts Journal*, pp. 8–20, 1995.

[8] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence properties of the nelder–mead simplex method in low dimensions," *SIAM Journal on Optimization*, vol. 9, no. 1, pp. 112–147, 1998.

[9] P. Whitle, *Hypothesis testing in time series analysis*. Almqvist & Wiksells, 1951, vol. 4.

[10] L. D. Majmin, "The implications of investor behaviour to financial markets," Ph.D. dissertation, Imperial College London, 2012.

[11] P. C. Andreou, C. Charalambous, and S. H. Martzoukos, "Pricing and trading european options by combining artificial neural networks and parametric models with implied parameters," *European Journal of Operational Research*, vol. 185, no. 3, pp. 1415–1433, 2008.

[12] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the marquardt algorithm," *Neural Networks, IEEE Transactions on*, vol. 5, no. 6, pp. 989–993, 1994.