Comparison of Least Square Monte Carlo Algorithm and Binomial Tree Model for Pricing American Options

Yuxin Meng

New York University, New York, 10003, the United States

ym1933@nyu.edu

Abstract. The financial market is a rapidly growing industry where derivatives are popular among investors. The rapid growth of options trading leads to the development of various option pricing theories and models. Through them, Longstaff and Schwartz improved the Monte Carlo model in 2001. The improved least square Monte Carlo simulation (LSM) is widely used in pricing American options. This paper aims to compare two estimation methods in pricing American options, namely the Least Square Monte Carlo Algorithm and the Binomial Tree Model, and detect which model better estimates the accuracy of the operation. In a refinement of using the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model to measure the volatility, an empirical comparison of the models using the Copper future contract is conducted. It is found that the binomial tree method can more accurately predict the American options prediction problem, and applying it to the pricing of copper can not only improve the market but also provide a more reasonable and rapid pricing basis for its subsequent development.

Keywords: American Options, Least Square Monte Carlo, Binomial Tree, GARCH Model.

1 Introduction

In 1973, Fischer Black and Myron Scholes developed the Black-Scholes model (B-S model), which employs partial differential equations to estimate the price of the options. This method is widely used in option pricing because of its simplicity and clarity [1]. However, in the research of American option pricing, because of the possibility of the option to be exercised in advance, the pricing of an American option is more complex. Hence, the B-S model cannot accurately calculate American options. Therefore, in 2001, Longstaff and Schwartz proposed the least square Monte Carlo simulation to help price American options. The primary method for pricing complex options is LSM because it is easy to simulate [2]. Most of the current research directions are on how to use the least squares Monte Carlo method to estimate American options, but few of them focus on comparing the Monte Carlo simulation method to other previous American option pricing models. LSM has its own limitations such as the payoff function cannot be fully represented by the finite basis functions [3]. This paper aims to compare

H. Mallick et al. (Eds.): ICEMCI 2022, AEBMR 231, pp. 2019–2028, 2023.
https://doi.org/10.2991/978-94-6463-098-5_227

the applied binary tree with the pricing simulated by the least squares Monte Carlo simulation method, thus reporting a more accurate and efficient method. In the empirical process, the author uses the GARCH model to estimate the volatility of the model. The volatility estimated is applied to the LSM pricing simulation to calculate the copper option price to make the price simulation analysis. The same data is used in the binomial tree model to calculate the price. Finally, by comparing the error size of the two models, a more accurate model can be obtained. The paper is organized as follows: section 2 is devoted to the GARCH model, the LSM model, the binomial tree model, and the data sources. Section 3 shows the results of the predicted price and conducts a comparative analysis between the two models. This paper can offer some help to other scholars in option pricing research, provide pricing references for enterprises or individuals engaged in options trading, and promote the development of the copper market in a better direction.

2 Methodology

2.1 Data

In this paper, the data of Shanghai copper futures were chosen. China, Europe, and America are all major countries in copper consumption. China accounts for half of the copper consumption, therefore the demand for copper plays an essential role in the price of Shanghai copper futures. To obtain more reasonable and fully represented market price data, the CU2210 copper futures contract was selected for empirical analysis. The data were downloaded from the Tushare website. According to the CU2210 future contract, the daily closing price was selected from October 18, 2021 to August 26, 2022, for a total of 211 groups of data. By using Python, Figure 1 is obtained.

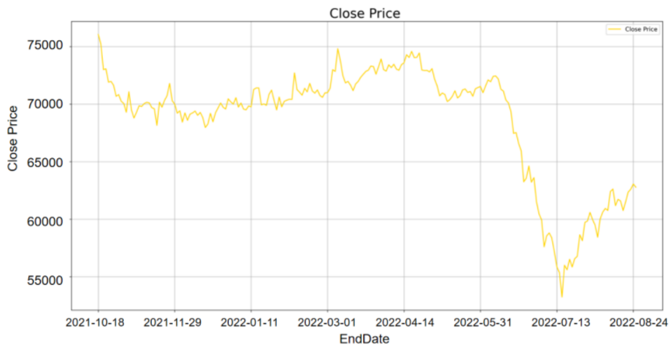

Fig. 1. Daily closing price chart of the copper futures contract (original).

2.2 GARCH Model

Volatility estimation is very important in option pricing. The accuracy of volatility estimation will directly affect the accuracy of subsequent pricing. Previously, the volatility estimation is using historical volatility. However, historical volatility has a significant lag in reflecting the rate of return. So, this paper will use the GARCH model to give a more accurate estimation of volatility. The GARCH model ensures that the return series is normally and conditionally distributed, and the error distribution is unsymmetric [4]. In the GARCH Model, the author assumes that the asset price is a discrete-time stochastic process. Duan [5] proposed that the returns of the asset follow a conditional lognormal distribution under the physical measure P as

$$ln\frac{X_t}{X_{t-1}} = r - \frac{1}{2}h_t + \lambda_1\sqrt{h_t} + \varepsilon_t \tag{1}$$

where the asset price at time t is denoted as X_t, r is the risk-free interest rate, λ_1 is the asset premium, and ε_t follows a GARCH(p,q) process following a normal distribution with mean zero and conditional variance h_t. And

$$h_t = \alpha_0 + \sum_{i=1}^{q}\alpha_i\,\varepsilon_{t-i}^2 + \sum_{j=1}^{p}\beta_j h_{t-j} \tag{2}$$

where $\alpha_0 \geq 0, \alpha_i \geq 0$ for $i = 1,2,\dots,q$ and $\beta_j \geq 0$ for $j = 1,2,\dots,q$.

In this paper, the author only focuses on the GARCH(1,1) case, so Equation(2) is simplified to

$$h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1 h_{t-1} \tag{3}$$

Here, $\alpha_1 + \beta_1 < 1$ suffices for wide-sense stationarity.

2.3 Binomial Tree Models

An n-period binomial tree is a stochastic model estimating the dynamics of a stock price changing through time [6]. First, the whole option time is divided to get a very small period Δt. Then according to the no-arbitrage theory, each point price will rise and fall, and over time, the path of the price will grow exponentially, which forms a binary tree map of the stock price movement until the option deadline. The option price of the previous node is then calculated forward from the price of the last time node (the price of time T), all the way until the present moment, so as to get the option prices[7]. More specifically, the steps are the following:

Step 1: Creating the Binomial Price Tree. The underlying price only has two trends: going up and down. It is a discrete-time stochastic process that

$$S^{(n)}(i+1) = \begin{cases} uS^{(n)}(i), & \text{if prices increase from } i \text{ to } i+1 \\ dS^{(n)}(i), & \text{if prices decrease from } i \text{ to } i+1 \end{cases} \tag{4}$$

where it is required that $u > d$ and $S^{(n)}(0) = s$. And by the Cox, Ross, and Rubinstein (CRR tree) model, the up and down factors are calculated using volatility, and the time duration of a step is measured in years [1]. Upward or downward tendencies in the log prices are incorporated in the model, which is determined by the parameter specifications:

$$u = e^{\sigma\sqrt{\Delta t}} \quad (5)$$

$$d = e^{-\sigma\sqrt{\Delta t}} = \frac{1}{u} \quad (6)$$

Step 2: Finding the Option Value at the Final Node. After learning u and d, the price of the asset at each node can be calculated. Then, there is a need to decide the option value at the final node (the expiration date of the option – time T). The option value is:

$$MAX\,[(S_n - K), 0] \;\; for\ call\ options$$

$$MAX\,[(K - S_n), 0] \;\; for\ put\ options$$

where K is the Strike price, and S_n is the price of the underlying asset at the nth period (expiration date).

Step 3: Calculating Option Value at Previous Nodes. Now there is a need to discount the value at time T to the previous node T-1. According to the Risk-Neutral Valuation and Replication, the expected value is calculated using the option values from the later two nodes weighted by their respective probabilities, that is:

$$V_1 = [pS_u + (1 - p)S_d]e^{-r\Delta t} \quad (7)$$

where r is the risk-free interest rate and p is the probability of an up move in the underlying, $(1 - p)$ is the probability of a down move. And p is chosen so that the related binomial distribution can simulate the geometric Brownian motion of the stock:

$$p = \frac{e^{r\Delta t} - d}{u - d} \quad (8)$$

Then the formula in step 2 is used to calculate the value at the T-1 node, named V_2. Comparing V_1 and V_2, the option is exercised if V1 is higher, and otherwise, it is not. Keep repeating the process, and the value at all previous nodes can be calculated, so as to decide whether there is a need to exercise or not.

2.4 Least Square Monte Carlo Simulation (LSM)

It is computationally infeasible to handle more than a couple of stochastic factors in the binomial tree model because the number of nodes will grow exponentially in the number of factors as mentioned in 2.3 [8]. Another method is to use simulations. In 2001, Longstaff and Schwartz proposed the Least Square Monte Carlo approach, which is to estimate the conditional expectation from the cross-sectional information in the simulation by using the least squares [1]. In general, LSM is just the combination of the Monte Carlo method and the nonlinear regression. Least-squares method is a way to resolve the regression analysis. For the basic idea, the least-squares method is applied to estimate α parameters at model $Y = \alpha X + \varepsilon$ [9], which give minimum the sum of squared errors as.

$$S(\alpha\alpha) = (Y - X\alpha)'(Y - X\alpha) \tag{9}$$

LSM will find the optimal stock price for the American option to execute the options , which is to consider the intersection between the line function and the payoff function [9]. The idea is as follows:

There is an underlying complete probability space (Ω, F, P) and a finite time horizon [0, T] where the state of Ω is the set of all possible realizations of the stochastic economy between time 0 and T, F represents the sigma filed of all possible events at time T, and P is the probability measure defined on the elements of F [1]. And $F = \{\mathcal{F}_t; t \in [0, T]\}$. Assume $\mathcal{F}_t = F$ in this case, and the existence of an equivalent martingale measure Q for this economy.

F can decide the best exercise node at each simulation path, where the value of an American option equals the maximized value of the discounted cash flow. Then, the notation $C(\omega, s; t, T)$ denotes the possible cash flows of the simulated paths generated by exercising the options [10]. Assume the best exercising time is s, then the option holder will own the option from time t to s, and $t \leq s \leq T$.

This paper only focuses on the case where the American can only be exercised at discrete times. Assume there are K time points, which is $0 < t_1 < t_2 < \cdots < t_k = T$. In order to approximate the price of these options, K is taken sufficiently large. And at time t_k, the option holder can compare the immediate exercise value to the value of continuation. The immediate exercise price equals the cash flow at the time t_k and the value of continuation is taken by taking the expectation of the remaining discounted cash flows $C(\omega, s; t_k, T)$ concerning the risk-neutral pricing measure Q. That is

$$F\ (\omega; t_k) = E_Q\left[\sum_{j=k+1}^{K} exp\left(-\int_{t_k}^{t_j} r(\omega, s)ds\right) C\left(\omega, t_j; t_k, T\right) \mid \mathcal{F}_{t_k}\right] \tag{10}$$

where ω is the number of simulated paths, and $r(\omega, s)$ is the riskless discount rate. Then LSM uses least squares to approximate the conditional expectation function at $t_{k-1}, t_{k-2}, \dots, t_1$. And to make variables unrelated to each other, this model employs the Laguerre formula to establish an orthonormal basis. Assume that X is the value of the asset underlying the option and that X follows a Markov process. With these specifications, $F\ (\omega; t_{k-1})$ can be represented as

$$F\ (\omega; t_k) = \sum_{j=0}^{\infty} a_j L_j(X) \tag{11}$$

where a_j coefficients are constants.

3 Results and Analysis

3.1 Estimation Volatility Using the GARCH Model

Through the introduction of the data and the model presented above, the GARCH (1,1) model is used first to estimate the daily volatility of the copper future. Equation (3) can be remade as

$$\sigma_t^2 = \gamma V_L + \alpha\mu_{t-1}^2 + \beta\sigma_{t-1}^2 \quad (12)$$

where V_L is the long-term variance, γ, α, β represent the weight, and $\gamma + \alpha + \beta = 1$. Then, the sequences of volatility can be built by using the GARCH (1,1) in python: by checking the official website of China Metal Market, it is known that copper futures options CU2210 still have 21 trading days before the final exercise date.

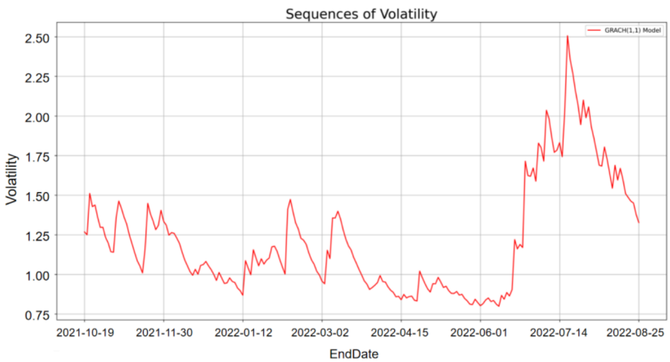

Fig. 2. Sequences of Volatility (original).

Therefore, in order to better use the LSM and binomial tree model to estimate the pricing of the copper options, the author first predicts the volatility of the copper futures returns for the next 21 days. According to the estimation obtained by Python, the following results are obtained:

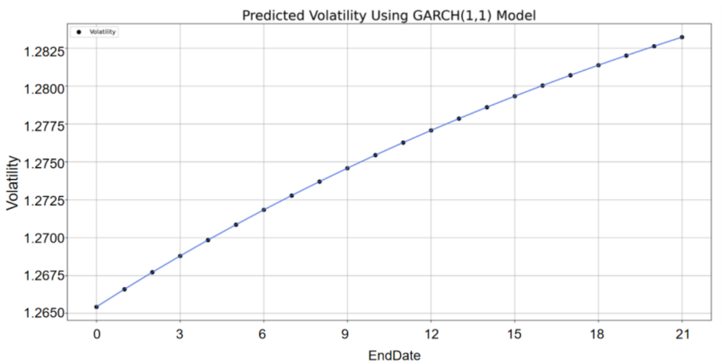

Fig. 3. The predicted volatility using the GARCH (1,1) model (original).

From Figure 3, it can be seen that the volatility keeps increasing but does not change much in the 21 days. And the data will be used later in the two models.

3.2 Comparing Option Pricing Using the Two Models

For the selection of the discount rate, this paper uses the one-year SHIBOR interest rate, which is 1.96%. First, in the LSM simulation to price copper options, the algorithm uses the GARCH (1,1) model and parameters S0 = 62780, r=0.0196, Number of path=1000, day=22; then, through Python, the data is shown in Table 1:

Table 1. LSM price at each node (original).

	PATH 1	PATH 2	PATH 3	…	PATH998	PATH 999	PATH1000
1	62780.00	62780.00	62780.00	…	62780.00	62780.00	62780.00
2	62762.89128	61692.92	63339.89	…	62840.2146	61970.8	62689.79
3	62747.257	62549.979	63503.593	…	63036.427	61119.502	63287.991
4	62834.493	63649.005	64792.656	…	61893.69726	58866.38	63073.369
…	…	…	…	…	…	…	…
21	63510.917	60771.617	69226.928	…	65675.616	563889.255	64311.933
22	65339.119	61518.93	67915.166	…	65790.08	56765.93	64894.82
23	65714.44	63129.836	67142.5	…	65954.834	56040.07	65586.735

It can be seen that due to the small path and the large time interval, the error of estimating the prices using the LSM simulation method is large. Hence, in order to make the price simulation more accurate, the author selects a large enough number of paths. By comparing the simulated exercise value of the simulated futures price of each node with the size of the continuation value, a judgment can be made on whether to continue to hold or directly exercise the option and take the larger price as the expected profits at the current node. Finally, the maximum return of each path can be obtained and discounted according to the risk-free return. Taking the average value, the call option simulation price at the current moment is 1563.59 RMB, and the put option price is 1659.49 RMB. Similarly, using the same option data to estimate the ending price in the binomial tree model, the call option price is 1426.97 RMB, and the put option price is 1543.21 RMB. Then the two models will be used to simulate active options on CU2210 futures contracts based on the market, and the comparison is shown in Table 2.

Table 2. Comparison of the ending prices of the two models (original).

	Close Price	ContractAbbr	TradingCode	StrikePirce	CRR_price	MC_price
1	1572	Oct. Copper put options 63000	CU2210P63000	63000	1543.213375	1638.049196
2	2120	Oct. Copper put options 64000	CU2210P64000	64000	2122.525112	2272.99869
3	2862	Oct. Copper put options 65000	CU2210P65000	65000	2802.708611	2959.883435
…	…	…	…	…	…	…
43	8718	Oct. Copper call options 54000	CU2210C54000	54000	8880.693862	8869.710574

44	7760	Oct. Copper call options 55000	CU2210C55000	55000	7890.987591	7905.908602
45	1378 8	Oct. Copper call options 49000	CU2210C49000	49000	13866.52522	13862.50299

Here the "ClosePrice" is the settlement of the options, "CRR_price" is the estimated price using the binomial tree model, and "MC_price" is the estimated price using LSM. There is a total of 45 options, and the picture shows 22 of them. From Table 2, it can be observed that, as the final exercise date of CU2210 copper futures options approaches, the simulated option price using LSM and binomial tree model gets closer to the real price, and their difference is not that huge. So, to estimate their accuracy, there is a need to compare their estimation errors using RMSRE (root mean square error) and AARE (average absolute relative error). Their formulas are:

$$AARE = \frac{1}{N}\sum_{i=1}^{N}\frac{|P_i^F - P_i^R|}{P_i^R} \tag{13}$$

$$RMSRE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{(P_i^F - P_i^R)^2}{(P_i^R)^2}} \tag{14}$$

where P_i^R is the real price, and P_i^F is the forecast price. Calculated by Python, the results are shown in Table 3:

Table 3. Estimation Error Comparison (original).

	AARE	RMSRE
CRR	0.330448255474434	0.330448255474434
MC	0.42981765253214	0.42981765253214

From Table 3, the estimation error for the binomial tree model is a little bit less than the least square Monte Carlo simulation. Therefore, by comparing the two models and estimating the same-day option price, a conclusion can be drawn that the binomial tree model is more accurate. However, in general the two models both show a relatively high error, so models can still be improved to better compare the errors.

4 Conclusion

With the rapid development of option derivatives, their role in the financial market is improving. The requirements on the option pricing model become higher. Therefore, choosing the right model to offer reasonable pricing is quite challenging. The first challenge is to select an appropriate volatility estimation model to estimate the volatility of the options of the assets. This paper uses the GARCH model to estimate volatility, with the advantage of calculating the volatility from discrete asset price data. Moreover, only a few parameters are needed to be estimated through the GARCH model. In order to achieve more precise pricing of options, this paper compares the LSM to the binomial tree model to simulate the pricing of copper futures and options. The results show that after using RMSRE and AARE to estimate the error, the binomial tree model is more accurate.

This paper has made innovations in research ideas and volatility selection. Moreover, through empirical analysis and comparison, the application of LSM and binomial tree model to Copper option pricing has made some contributions to the future pricing research of commodity options. However, there are still some aspects that can be further discussed and improved. First, this paper directly employs the one-year SHIBOR interest rate, which may not be able to accurately simulate the futures pricing. LSM can be more accurate when using the daily overnight interest rate. However, because the overnight interest rate changes every day, using the same interest rate may affect the pricing results. Future research can switch to linear interpolation of interest rates. Second, the GARCH model is used in the volatility estimates in this research, but because the GARCH model itself has some drawbacks, the model cannot reflect the asymmetrical characteristics of volatility. The positive and negative news will have different effects on the time-series volatility. Therefore, the GARCH model may not be the best model for the copper options . Future research can apply different estimation models for volatility and compare them to greatly reduce the error of the LSM method. Finally, this paper only incorporates the example of copper options, exceptions exist for the single comparisons. By including more examples of comparing models, the result can be more convincing.

References

1. Shinde, A. S., Takale, K. C.: Study of Black-Scholes Model and its Applications. Sciverse Science Direct (2012).
2. Jia, Q. Y.: Pricing American Options using Monte Carlo Methods. U.U.D.M. Project Report, 8 (2009).
3. Woo, J., Liu, C. R., Choi, J. H.: Leave-One-Out Least Square Monte Carlo Algorithm for Pricing American Options. Captital Markets: Asset Prcing & Valuation Journal (2018).
4. Ghahramani, M., Thavaneswaran, A.: A note on GARCH model identification. Sciverse Science Direct (2007).
5. Zhang, W. J., Zhang, J. E.: GARCH Option Pricing Models and the Variance Risk Premium. Journal of Risk and Financial Management (2020).
6. Korn, R., Müller, S.: Binomial Trees in option Pricing—History, Practical Applications and Recent Developments. In: Devroye, L., Karasözen, B., Kohler, M., Korn, R. (eds) Recent Developments in Applied Probability and Statistics. Physica-Verlag HD (2010). https://doi.org/10.1007/978-3-7908-2598-5_3.
7. Zhao, Q. B.: Empirical Study on American Futures option based on the Least Squares Monte Carlo Simulation—Take Sugar Futures option for example.
8. Stentoft, L.: Assessing the Least Squares Monte-Carlo Approach to American option Valuation. Review of Derivatives Research (7), 129–168 (2004).
9. Palupi, I., Sitorus, I. U., Umbara, R. F.: Using Lest-Square Monte Carlo Simulation to Price American Multi Underlying Stock Options. IEEE Explore, 2015.
10. Francsis, A., Schwartz, L. E. S.: Valuing American Options by Simulation: A simple Least-Squares Approach. The Review of Financial Studies Spring, 113-147 (2001).