Research on Minimum Variance Delta Hedging Strategy of CSI 300ETF Options

Xialin Du, Yajie Wang*
School of Management, Harbin Institute of Technology, Harbin, Heilongjiang, China
**Corresponding author.*

Abstract: This study constructs Hull & White and neural network-based minimum-variance Delta models, assessing Delta hedging strategies via the Gain metric. Using CSI 300ETF option data (2019/12/23–2023/12/31), we compare LSTM and CNN volatility forecasts and Delta hedging efficacy. Results show the LSTM-based strategy outperforms Black-Scholes and Hull & White approaches, offering optimal risk management for SSE 300ETF options. By integrating volatility dynamics and price reversals, the framework addresses traditional Delta-neutral hedging limitations, advancing practical risk mitigation methodologies.

Keywords: Delta Hedging Strategy; Minimum Variance Delta Hedging; Implied Volatility; Options on CSI 300ETF

1. Introduction

The rapid proliferation of financial derivatives, particularly options, has transformed portfolio construction and risk management practices. Since the introduction of SSE 50ETF options in 2015, China's options market has experienced exponential growth, with trading volume surging to 45.35 million contracts, up 185% year on year and cumulative turnover reaching ¥285.9 billion, up 67% year on year, by mid-2023. This growth underscores the market's pivotal role in China's financial ecosystem.

Among derivatives, ETF options—notably the CSI 300ETF option launched in late 2019—have gained prominence due to their liquidity, representativeness, and hedging utility. As a key instrument tied to the CSI 300ETF, this option reflects investor sentiment in A-shares and fund markets while enabling risk mitigation in spot portfolios. However, option pricing dynamics are inherently nonlinear, driven by multifactorial influences, complicating risk quantification. Traditional Delta-neutral strategies, reliant on Black-Scholes-derived Greeks, face limitations in minimizing portfolio variance, prompting innovations in minimum-variance Delta frameworks [1-3].

This study advances option risk management by integrating implied volatility forecasting and minimum-variance Delta hedging for CSI 300ETF options. First, we employ LSTM and CNN models [4] to capture nonparametric implied volatility patterns, enhancing predictive accuracy for market practitioners. Second, we benchmark Hull & White's quadratic model to derive and validate a minimum-variance Delta strategy tailored to China's options market. Finally, we compare parametric and nonparametric Delta hedging efficacy against Black-Scholes, identifying optimal risk mitigation approaches.

Theoretical contributions include refining implied volatility measurement via deep learning and expanding minimum-variance Delta theory [5]. Practically, this work equips investors with tools to align hedging strategies with market dynamics, reduce portfolio risk, and navigate A-share market expectations. Furthermore, our findings offer insights for developing derivative products and supporting China's real economy.

Existing research predominantly focuses on SSE 50ETF options, with limited exploration of CSI 300ETF instruments. While parametric volatility models are well-studied, nonparametric approaches using LSTM/CNN remain underexplored [6-8]. Similarly, despite global adoption of minimum-variance Delta strategies, their application in China is nascent. By addressing these gaps through deep learning-driven volatility modeling, this study provides novel methodologies for enhancing option risk management frameworks.

2 Model

2.1 Minimum Variance Delta

2.1.1 Minimum Variance Delta Hedging Principle

Delta-neutral strategy usually uses the Delta calculated by Black-Scholes as the measurement index. However, the assumption that the stock price obeys the geometric Brownian motion in the Black-Scholes formula is not consistent with the reality, and the assumption of constant volatility fails to depict the smile feature of the volatility curve. Due to the leverage effect and volatility feedback mechanism, there is an inverse relationship between the volatility and the underlying price, but the Black Scholes model does not correctly express this inverse relationship [9,10]. Black-Scholes Delta is not the optimal indicator for delta neutral hedging because it cannot be used to minimize the variance dispersion of the hedging position value [11,12]. The Delta indicator used for hedging should take into account both the change of the underlying asset price and the change of the volatility and the inverse relationship between the two.

As the defects of Black-Scholes are constantly discovered, the academic community has also explored a more appropriate Delta measure. The variance minimization Delta is a hedging measure that considers both the underlying stock price changes and the impact of expected volatility changes on the condition of the underlying stock price changes. The minimum variance Delta is the value of ($\Delta f - \delta^{mv}\Delta S$) with the smallest variance δ^{mv} , where Δf , ΔS represent the slight change of the option value and the underlying price respectively, and can also be expressed as the differential form of the corresponding parameter. Therefore, to calculate the variance minimization Delta, a model must be used to determine the expected change in volatility caused by the change in the underlying price and the expected change in the underlying price.

2.1.2 Implied Volatility and Minimum Variance Delta

As mentioned in 2.1.1, the minimum variance Delta is a hedging indicator that can take into account both the underlying price changes and the expected volatility changes, and it is the indicator σ_{imp} that reflects the expected volatility of options. Alexande et al. (2012) [13] found that the derivation process of the minimum variance Delta expression after volatility smile adjustment is as follows :

Considering only the influence of the underlying price S and implied volatility on the change of the option price σ_{imp}, we can obtain an option pricing formula based on the Black-Scholes formula and considering only the above two parameters. By Taylor expansion, we can obtain:

$$\Delta f = \frac{\partial f}{\partial S}\Delta S + \frac{\partial f}{\partial \sigma_{imp}}\Delta\sigma_{imp} + o\left(\Delta S^2\right) = \delta_{bs}\Delta S + v_{bs}\Delta\sigma_{imp} + \varepsilon \tag{1}$$

To get $\Delta f - \delta^{mv}\Delta S$, subtract $\delta_{bs}\Delta S$ both sides of (1) to get the equation of $\Delta f - \delta_{bs}\Delta S$ as:

$$\Delta f - \delta_{bs}\Delta S = (\delta_{bs} - \delta_{mv})\Delta S + v_{bs}\Delta\sigma_{imp} + \varepsilon \tag{2}$$

Taking the conditional expectation ΔS of both sides of (2), we can obtain:

$$\delta^{mv} = \delta_{bs} + v_{bs}\frac{E(\Delta\sigma_{imp})}{\Delta S} + \frac{E(\varepsilon)}{\Delta S} \tag{3}$$

In the case of infinitesimal ΔS, this term can be regarded as 0, and the last term can be regarded as infinitesimal, so Equation (3) is the calculation formula of the minimum variance Delta:ΔS

$$\delta^{mv} = \delta_{bs} + v_{bs}\frac{E(\Delta\sigma_{imp})}{\Delta S} \tag{4}$$

In order to find the minimum variance Delta and carry out subsequent hedging Gain comparison, it is necessary to find the expected implied volatility. The expected implied volatility must be predicted by modeling, so this paper chooses to use deep learning method to establish a nonparametric model for prediction.

2.1.3 Hull and White Minimum Variance Delta

The calculation formula of the minimum variance Delta is described in 2.3.2, but it is difficult to solve some parameters of $\frac{E(\Delta\sigma_{imp})}{\Delta S}$. Therefore, Hull and White proposed a quadratic model that can be used to calculate the minimum variance Delta through derivation.

Without considering the dividend, the vega of the European option can be expressed as. $v_{bs} = S\sqrt{T}G(\delta_{bs})$. Hull and White use the S&P 500 option-related data to estimate δ^{mv} in (4), and empirically find that the correlation of $\delta_{bs} - \delta_{mv}$ with the remaining maturity of the option is weak, and is a quadratic function of δ_{bs}. Therefore, Hull and White (2017) [14] proposed the quadratic function hypothesis about δ^{mv}. From the quadratic expression, we can obtain the relationship between the minimum variance Delta and the Black-Scholes Delta, where a, b and c are the parameters to be estimated:

$$\delta^{mv} = \delta_{bs} + \frac{v_{bs}}{S\sqrt{T}}\left(a + b\delta_{bs} + c\delta_{bs}^2\right) \quad (5)$$

Combining (4) and (5) yields the formula for the expectation of the change in implied volatility:

$$E\left(\Delta\sigma_{imp}\right) = \frac{\left(a+b\delta_{bs}+c\delta_{bs}^2\right)}{S\sqrt{T}}\frac{\Delta S}{S} \quad (6)$$

Based on equations (1), the formula used to perform parametric regression can be obtained:

$$\Delta f - \delta_{bs}\Delta S = \frac{v_{bs}}{\sqrt{T}}\left(a + b\delta_{bs} + c\delta_{bs}^2\right)\frac{\Delta S}{S} \quad (7)$$

In order to facilitate the calculation of subsequent Gain indicators, we also need to obtain the error terms under the Hull and White quadratic benchmark models, and the specific calculation formula is as follows:ε_{HW}

$$\varepsilon_{HW} = \Delta f - \delta_{bs}\Delta S - \frac{v_{bs}\Delta S}{S\sqrt{T}}\left(a + b\delta_{bs} + c\delta_{bs}^2\right) \quad (8)$$

Where, Δf represents the change of the corresponding option price of a certain variety on a certain day, ΔS represents the change of the underlying asset price, T represents the remaining $\delta_{bs} and\, v_{bs}$ are respectively the Delta and Vega values calculated under the Black-Scholes formula, which can be directly observed. And are the delta and vega values calculated under the Black-Scholes formula. The parameters a, b and c of the regression equation of the whole data set can be obtained by the relevant processing of the daily trading data of the CSI 300ETF options by (5).

2.1.4 Minimum Variance Delta Under Non-Parametric Model

Based on formula (4), we can obtain the most critical value of δ^{mv} in the general case. In addition, due to the form of implied volatility expectation $E\left(\sigma_{imp}\right)$ and the number of options products corresponding to different varieties and maturities of CSI 300ETF in the sample data, this paper improves to calculate the corresponding value of daily minimum variance Delta by output the predicted value one by one and then taking the average value. The specific calculation formula of minimum variance Delta δ_{DL}^{mv} under deep learning is as follows:

$$\delta_{DL}^{mv} = \delta_{bs} + v_{bs}\frac{E(\sigma_{imp})}{\Delta S} = \delta_{bs} + v_{bs}\frac{\sum_{i=1}^{n}\sigma_i^{imp}/n}{\Delta S} \quad (9)$$

In order to facilitate the calculation of subsequent Gain indicators, we also need to obtain the error term under the deep learning model, and the specific calculation formula is as follows:ε_{DL}

$$\varepsilon_{DL} = \Delta f - \delta_{DL}^{mv}\Delta S \quad (10)$$

Where, Δf represents the change of the corresponding option price of a certain variety on a certain day, and ΔS represents the change of the price of the underlying asset.

2.2 Gain Index

This paper needs to evaluate the effectiveness of the minimum variance Delta hedging under the parametric model and the nonparametric model and compares them with the Black-Scholes Delta hedging results respectively, and finally obtains the Delta metric that is more suitable for the delta-neutral hedging of CSI 300ETF options. Referring to the formula of Hull and White (2017) for the evaluation index of Delta hedging returns, the Gain index is established:

$$Gain = 1 - \frac{SSE(\varepsilon_{mv})}{SSE(\varepsilon_{bs})} \quad (11)$$

Where, SSE represents the sum of squared residuals, can be specifically written as:ε_{bs}

$$\varepsilon_{bs} = \Delta f - \delta_{bs}\Delta S \quad (12)$$

In this paper, the Black-Scholes Delta return is used as the unified measurement standard, and the minimum variance Delta return under the better deep learning model and the minimum variance Delta return of Hull and White's quadratic model are respectively selected to compare with BS Delta return. When the Gain is 0, it indicates that the minimum variance Delta and Black-Scholes Delta hedging strategies have the same effectiveness, and both strategies can be used in real trading. When the Gain is less than 0, the sum of squared residuals of the Black-Scholes Delta in the denominator is smaller, and the hedging strategy has better returns and better effects. When Gain is greater than 0, the sum of squared residuals of the minimum variance Delta in the numerator is smaller, and the hedging strategy has better returns and better effects.

3. Empirical Analysis

3.1 Data Selection and Descriptive Statistics

3.1.1 Data

As options in China are still in their infancy, and the CSI 300 ETF options were listed relatively late, the empirical analysis in this paper selects trading day data from the listing date, i.e., December 23, 2019, to December 31, 2023. Additionally, since volatility modeling requires data on the risk-free rate and the underlying asset, we include the CSI 300 ETF (code: 510300) over the same period, resulting in 977 trading days and 113,592 data entries.

The daily trading data for CSI 300 ETF options and the CSI 300 ETF fund are sourced from the Wind Financial Terminal. The option data includes trading date, contract code, expiration date, option type, strike price, open price, close price, high price, low price, Delta, Gamma, Vega, Theta, and trading volume. The ETF fund data includes trading date and close price.

The Shanghai Interbank Offered Rate (SHIBOR) has been widely adopted in academic research as a proxy for the risk-free rate due to its derivation from high-credit bank quotes with negligible default risk [15-16]. Following Wu Xinyu et al., this study uses SHIBOR as the risk-free rate, with data obtained from the China Money Network.

3.1.2 Sample Selection and Data Processing

To exclude extreme values and noise, we apply the data filtering methods of Chen Rong and Zhao Yongjie (2017) and Zheng Zhenlong and Wang Raosixing (2017) to the CSI 300 ETF options:

① Remove options with less than 7 days or more than 180 days to expiration to avoid liquidity and maturity effects. ② Exclude daily trading volumes below 10 to ensure price discovery validity. ③ Eliminate data violating theoretical price boundaries ④ Remove days with missing data for options, the ETF, or SHIBOR. After filtering, 83,625 option entries remain. Annualized time-to-maturity is calculated and added as a new column.

3.1.3 Descriptive Statistics

Figure 1 shows the closing price trend of the CSI 300 ETF from December 23, 2019, to December 31, 2023. The closing price fluctuated between 3.369 and 5.807 CNY, with a mean of 4.395 CNY. Prices generally rose until February 10, 2021, followed by a decline to around 3.5 CNY.

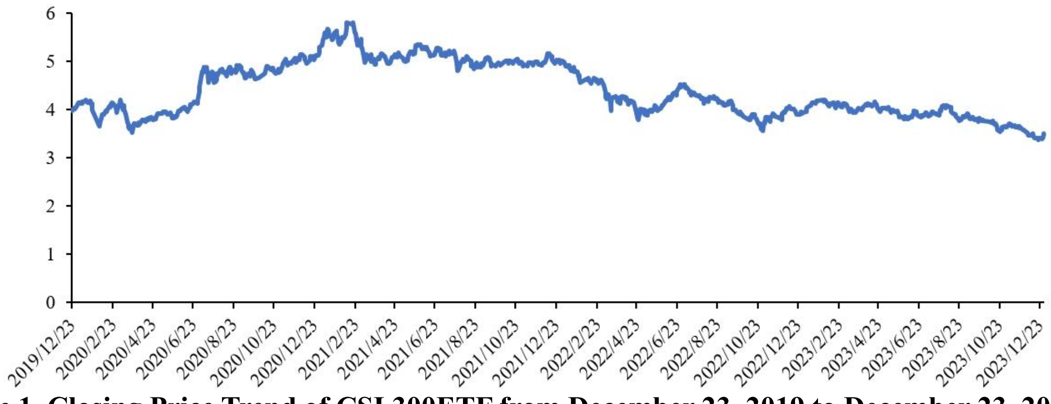

Figure 1. Closing Price Trend of CSI 300ETF from December 23, 2019 to December 23, 2023

Table 1 shows the descriptive statistics of the variables of the full sample in this paper. It can be seen that the mean of each variable is greater than the median, indicating that the distribution of sample variable data has obvious right-skewed characteristics. Among them, the mean and median values of daily strike price and Greek letters Gamma, Vega and Theta are similar, while the mean and median values of trading volume differ greatly, and the average value is significantly more affected by the maximum value than the median value. From the Delta data, we can see that the value of 50% is less than 0, which may be caused by the large number of put options in the whole sample. It can be seen from the price-related data that the option price of CSI 300ETF is basically between 0 and 1.9, and the median of all price data is less than 0.2. The maximum value of the annualized option remaining term T is 0.485, indicating that the maximum number of remaining days in the cleaned sample is 175 days and the minimum number is 8 days. In addition, the mean, median and standard deviation of implied volatility analyzed in the following section are 0.226, 0.153 and 0.092, respectively. The dispersion degree is smaller than that of other trading day price variables.

Table 1. Descriptive Statistics of Variables in the Full Sample

Model type	Mean value	Median	Maximum	Minimum	Standard deviation
Option strike price	4.431	4.4	6.750	3.000	0.752
Option open price	0.262	0.144	1.911	0.000	0.307
Option closing price	0.262	0.142	1.901	0.000	0.309
The option's highest intraday price	0.279	0.162	1.916	0.000	0.317
Option low intraday price	0.246	0.127	1.901	0.000	0.298
Options volume	16181.830	2660	446583	11	42288.430
Option remaining maturity T	0.194	0.153	0.485	0.022	0.130
Implied volatility	0.226	0.204	1.542	0.043	0.092
Delta	0.035	-0.04	1.000	-1.000	0.607
Gamma	0.654	0.536	4.743	0.000	0.543
Vega	0.004	0.004	0.015	0.000	0.003
Theta	-0.001	-0.001	0.001	-0.011	0.001

3.1.4 Sub-sample Statistics of Implied Volatility

According to the nature of options, there are 41252 call options and 42373 put options in the sample, and their implied volatilities are descriptively analyzed. It can be seen from Table 2 that the mean and median of implied volatility of call options are 0.200 and 0.186, respectively, which are lower than 0.251 and 0.230 of put options. The range of volatility of call options is 1.356, which is also lower than 1.481 of put options. The standard deviation of put options is 0.101, while that of call options is 0.073. That puts the volatility of discrete degree is greater than the call option. Call options implied volatility, in general, the overall value and volatility change were significantly less than the put option, that Shanghai and Shenzhen 300 ETF traders expect the future Shanghai and Shenzhen 300 ETF fund prices rose less likely in the future, for the call option price fluctuations have less expectation speculative strong put options.

From the perspective of intrinsic value, there are 40646 options with real value, 42886 options with dummy value and 93 options with value option in Table 3. Intrinsic value refers to the present value of the return that can be obtained when the option is exercised immediately. If the return is greater than 0, it is the real value option, otherwise it is the dummy value option, and equal to 0 is the flat value option. Table 3 shows that average value and virtual value options solid basic same, keep three decimal places are 0.226, flat value options of averages of 0.202, less than virtual value and real value of the option. The range of the three options is 1.5, 0.318 and 1.412, respectively. It can be seen that the maximum and minimum values of the options are very close, and the range is much smaller than that of the options with real value and real value. The standard deviations of the three options are similar to the range. The maximum of the real option is 0.107, the second is 0.074 for the dummy option, and the minimum is 0.056 for the even option, indicating that the volatility dispersion of the real option is the largest, and the volatility distribution of the even option is more concentrated, and the variation range is the smallest.

Table 2. Implied Volatility Statistics of Options with Different Attributes

	Call options	Put options
Number of samples	41252	42373
Average	0.200	0.251
Median	0.186	0.230
Maximum	1.399	1.543
Minimum	0.043	0.062
Standard deviation	0.073	0.101

Table 3. Implied Volatility Statistics of Options with Different Intrinsic Values

	At-the-money options	Even value option	Out-of-the-money options
Number of samples	40646	93	42886
Average	0.226	0.202	0.226
Median	0.197	0.186	0.210
Maximum	1.543	0.430	1.486

Minimum	0.043	0.112	0.074
Standard deviation	0.107	0.056	0.074

According to the remaining maturity of options, there are 19366 short-term options, 36518 medium-term options and 27741 long-term options in the sample. It can be seen from Table 4 that the maximum average value of short-term options is 0.265, the average value of medium-term options and long-term options are 0.219 and 0.207 respectively, and the medians of the three are very close to each other, all of which are around 0.2. From the analysis of dispersion degree, the range and standard deviation of the short-term option are the largest, which are 1.48 and 0.135 respectively. The medium-term options are 0.097 and 0.075, respectively. The range and standard deviation of the long-term option are the smallest, which are 0.645 and 0.06 respectively. All data, clear the duration of the option characteristics in accordance with our perception of implied volatility smile curve characteristics, the expiring options, the greater the volatility of any changes as the target price can affect the holder of the benefit in the end. However, the time from the final delivery date of long-term options is very long, and the impact of the underlying price on the final outcome is not obvious. Therefore, the market pays little attention to the option, and the implied volatility value and variation range are very small.

Table 4. Implied Volatility Statistics of Options with Different Remaining Maturities

	Short term options	Medium-term options	Long-term options
Number of samples	19366	36518	27741
Average	0.265	0.219	0.207
Median	0.225	0.202	0.197
Maximum value	1.542	1.040	0.693
Minimum	0.062	0.043	0.048
Standard deviation	0.135	0.075	0.060

3.1.5 Analysis of Implied Volatility Fluctuations

This section examines the characteristics of implied volatility dynamics. Given the extensive variety of option contracts, two specific contracts:510300C2009M03900(call option) and 510300P2009M03900(put option) with identical expiration dates (September 2020) and strike prices (3.9 CNY) are selected for comparative analysis. Both contracts are linked to the CSI 300 Exchange-Traded Fund (ETF) (code: 510300).

As depicted in Figure 2, the implied volatility of both contracts demonstrates a volatility smile pattern, characterized by an initial decline followed by an upward trajectory as expiration approaches. This reflects market participants' heightened sensitivity to price fluctuations during the early and late phases of the contract lifecycle, with reduced emphasis during intermediate periods. Key observations include: ① Call Option. The implied volatility curve exhibits relative smoothness, with a peak-to-trough differential of 0.26. Implied volatility reached its lowest level approximately three months after the contract listing and gradually increased as expiration neared. ② Put Option. Implied volatility displays pronounced variability, with a peak-to-trough differential of 0.62. Significant fluctuations occurred near expiration, including two localized spikes in July 2020.

The July 2020 volatility surge coincided with a sharp price increase in the CSI 300 ETF to 4.7 CNY on July 1, which rendered the put option deeply out-of-the-money (OTM). This price movement amplified market expectations of a potential future decline in the ETF's value, thereby driving a rapid rise in the implied volatility of the put option.

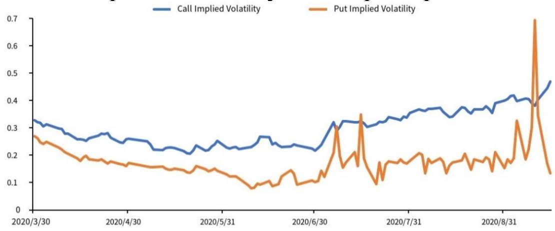

Figure 2. 510300 c2009m03900, 510300 p2009m03900 Options Implied Volatility

We here interpret that mechanisms underlying divergent Implied Volatility Behavior ① Asymmetric Risk-Return Structure. The value of a put option increases substantially if the underlying asset price declines but may expire worthlessly if prices stabilize or rise. This binary payoff structure intensifies investor focus on implied volatility dynamics as expiration approaches. ② Supply-Demand Dynamics. A limited supply of put options combined with heightened demand can accelerate price appreciation, further amplifying implied volatility.

3.2 Empirical Analysis of Non-Parametric Implied Volatility Forecasting Models

In the non-parametric modeling part, two deep models, LSTM and CNN, are used respectively, which are completed by Python 3.8.8. In the part of data cleaning and processing, pandas and numpy libraries are used. The two deep models are trained using Keras and sklearn libraries of Tensorflow.

3.2.1 LSTM Neural Network Model Training and Empirical Study

The daily trading data of all types of Shanghai-Shenzhen 300ETF options from December 23, 2019, to December 31, 2023 are used as the data set and the training set and test set are divided accordingly. 80% of all data is selected as the model training set and the remaining 20% as the test set. The rolling window method is used to generate the sample set, that is, the implied volatility of the next trading day is predicted by using the data of 10 trading days as input. In the process of LSTM model training, it is also necessary to convert the data of training set and test set into 3D tensor respectively.

In the process of model training, eight time series factor data of CSI 300ETF fund closing price, strike price, option remaining term, risk-free interest rate, the highest price, lowest price, opening price and closing price of each transaction of CSI 300ETF option are selected as input features.

In the model hyperparameter optimization setting, under the condition that the batch number of samples is 60 and the iteration number epoch is 100, the hyperparameter learning rate and time step size of the model are adjusted to optimize the model training. The value range of learning rate is 0.01, 0.001, 0.0001, and the time step is set between 20 and 140. The loss of LSTM model is investigated every 20 steps, and the loss of LSTM model under the combination of 20 groups of learning rate and time step is compared, and the optimal LSTM implied volatility prediction model is found accordingly.

As shown in Figure 3, when the learning rate is set to 0.01, with the increase of the number of time steps, the validation loss of the model shows an overall downward trend. Specifically, when the time step increases from 20 to 140, the model validation loss gradually decreases from 0.0044 to 0.0039, indicating that the model considering 140 trading days implied volatility data has the best prediction effect. Adding historical implied volatility data helps to improve the prediction performance of the LSTM model. From the figure, we can also find that when the learning rate is 0.001, the overall validation loss value is lower, and the same results are similar to the above, that is, the increase of time step helps to reduce the prediction loss value and improve the prediction performance of the LSTM model. At the same time, the paper also finds that the smaller learning rate makes the model more cautious when updating the weights, and the model can achieve a lower validation loss at an earlier time step under the learning rate, avoiding the oscillation near the optimal solution. When the learning rate is further reduced to 0.0001, the trend of validation loss is similar to the previous two conclusions, but the overall loss value fluctuates, which indicates that too small learning rate leads to too small step when the model weight is updated, which leads to the interference of noise data in the volatility information, and the convergence speed slows down.

In summary, the loss of LSTM model decreases rapidly at the initial stage of training, and finally reaches around 0.004, which converges to a good value, and the model prediction effect is good. In the training process, when the learning rate of the model is set at 0.001 and the step size is set at 140, the model has the best prediction performance for the implied volatility of CSI 300ETF options.

3.2.2 Convolutional Neural Network Model Training and Empirical Study

In the training process of CNN convolutional neural network, in order to effectively compare with LSTM, the same features as CNN are selected as input variables, which are respectively the closing price, strike price, remaining term of options, riskless interest rate, the highest price, lowest price, opening price and closing price of each transaction of CSI 300ETF options.

Since each feature is time series data, the model training is carried out according to the hyperparameter setting at the time of establishment, in which the number of time steps is 10. In addition, the number of samples used in one training and the number of iterations have a certain impact on the training efficiency and results of the model. In order to compare with LSTM neural network model, batch size is still selected as 60 and epoch is 100. In the training process, Reduce LR On Plateau callback function is used to reduce the learning rate, and its scaling value is set to 0.95, and the lower limit of learning rate is still 0.0001.

In Figure 4, it can be seen that the training loss decreases rapidly from the initial stage of training and gradually becomes stable after the number of iterations reaches 20, and converges to a smaller number. However, although the validation loss value is also relatively small, it is much higher than the training loss as a whole and shows an obvious upward trend, indicating that the CNN convolutional neural network prediction model has overfitting phenomenon, which performs well on the training set, but performs poorly on the test set data, and the generalization ability of the model is insufficient.

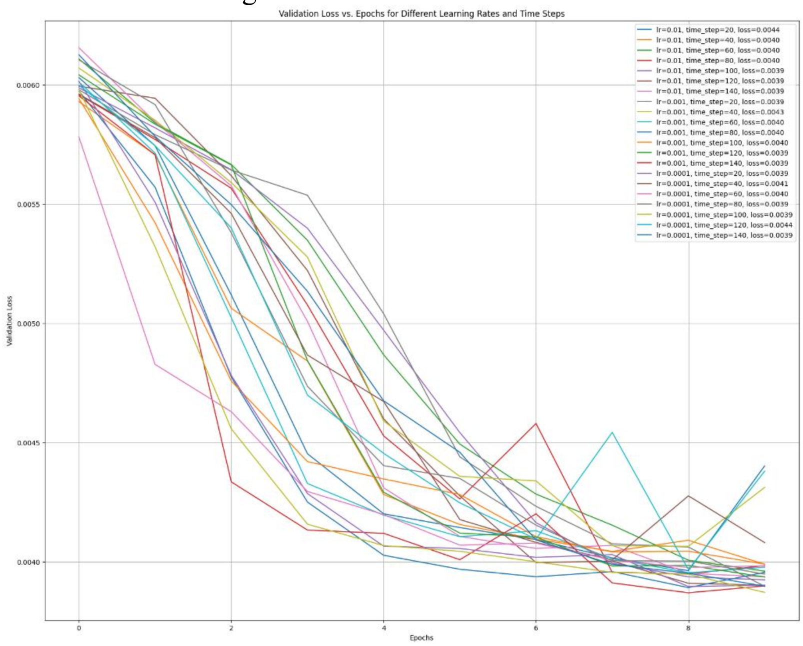

Figure 3. Validation Loss Relation of LSTM Model with Different Learning Rate and Step Size

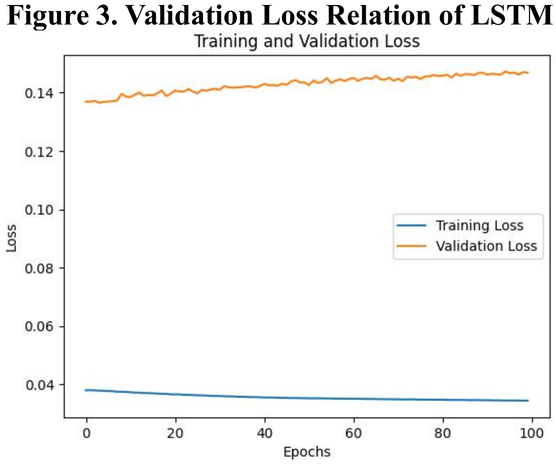

Figure 4. Loss of Convolutional Neural Network Implied Volatility Prediction Model

3.2.3 Comparison and Analysis of Model Results

In Figure 5 and Figure 6, there are about 17,500 implied volatility data in the test set, whose range is basically between 0 and 1.4, and the predicted data range is between 0.1 and 1.2 by LSTM model. The predicted trend is basically consistent with the change trend of the real data, except for some abnormal peaks and low extreme values, which can basically correctly describe the trend of implied volatility. However, its variation range does not cover the implied volatility of 0-0.1 part, so the predictive ability of LSTM model for real volatility needs to be strengthened.

In contrast, the predictive value range of CNN model has been expanded to -0.5 to 1.5, which is beyond the actual range of implied volatility, and the variation range of the predicted data of the training set is significantly higher than that of the true implied volatility. The fluctuation range of the predicted data is drastic, which is seriously out of line with the real market volatility, and cannot correctly reflect the variation of the market implied volatility. It shows the phenomenon of over-fitting, and the overall forecasting ability of the model is poor. Therefore, the LSTM model has better fitting effect than the CNN model, and is more suitable for the implied volatility modeling. It is also necessary to consider the specific performance index data for model selection, so the performance index of the two training models is described and evaluated below.

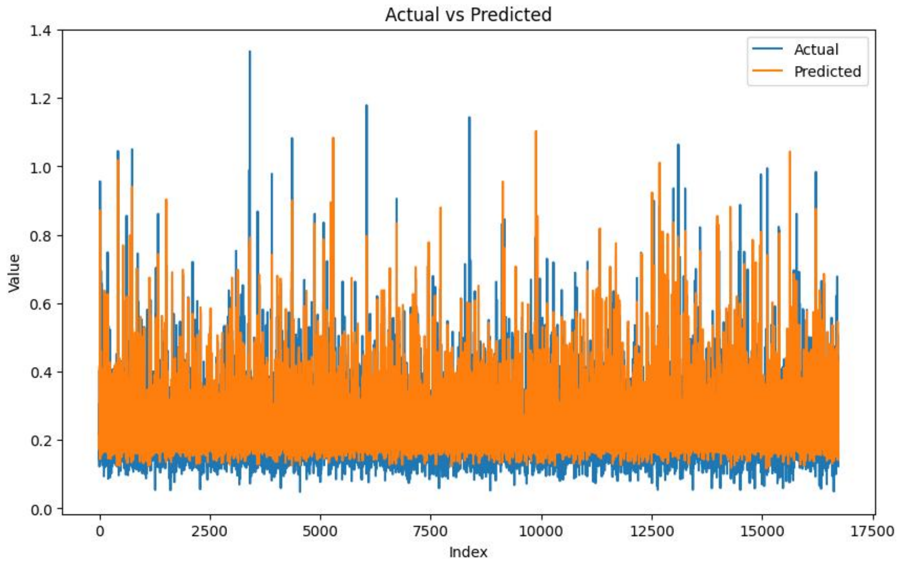

Figure 5. Comparison of Real and Predicted Values of LSTM Neural Network

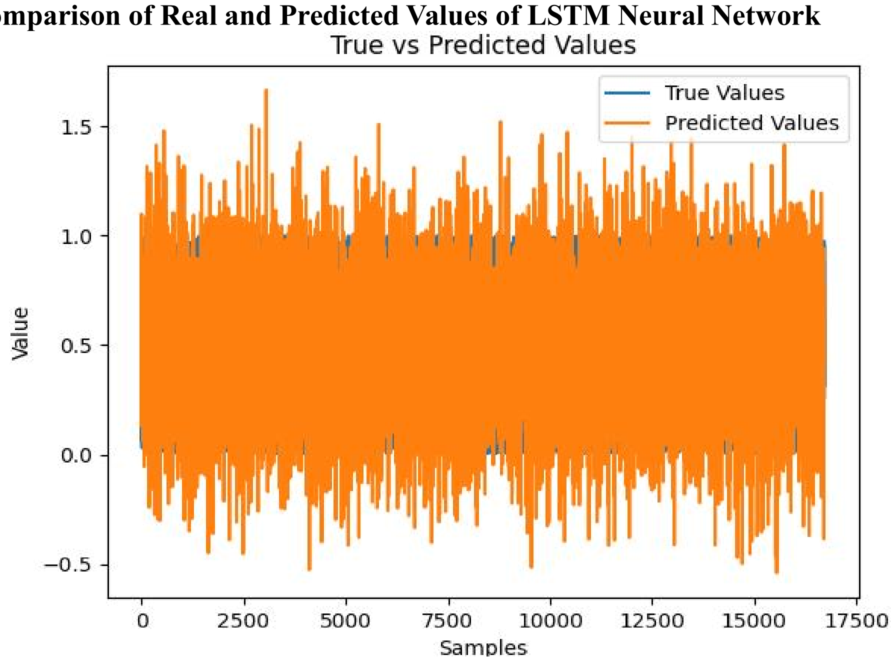

Figure 6. Comparison of Real and Predicted Values of Convolutional Neural Network

Meanwhile, Table 5 shows the performance comparison of two non-parametric models based on LSTM neural network and CNN convolutional neural network in the prediction of CSI 300ETF option implied volatility test set:

Table 5 The Degree of Fit and Prediction Error of LSTM and CNN Models

Model type	MSE	RMSE	MAE	MAPE
LSTM neural network	0.001	0.031	0.019	0.194
CNN convolutional neural network	0.147	0.383	0.312	455.523

Note: The smaller the annotation error value is, the better the model regression and prediction ability is

In Table 5 the MSE, RMSE and MAE of the LSTM neural network model are 0.001, 0.031 and 0.019 respectively, and the overall fitting and prediction effect is good. However, the CNN convolutional neural network was 0.147, 0.383, 0.312, respectively, and the error was significantly greater than that of the LSTM model. In addition, the MAPE, which represents the average absolute percentage error, is 19.4% and 45500% respectively, indicating that the training and prediction error of LSTM model is within 19.4%, while that of CNN convolutional neural network is very large, which is similar to the conclusion of CNN loss map, and its model has over-fitting phenomenon. The model has an over-fitting phenomenon. It has weak generalization ability for the prediction of CSI 300ETF options.

To sum up, the LSTM neural network model performs better than the other two nonparametric models, and its capture of implied volatility changes is more consistent with the real situation, showing its powerful regression and prediction ability for time series data. However, CNN convolutional neural network has an overfitting phenomenon, which indicates that it is not suitable to use this model to model the option volatility of Shanghai and Shenzhen 300ETF. At the same time, the training process of the LSTM neural network model also shows that eight time series factors, such as closing price, strike price, remaining term of the option and risk-free interest rate, have an impact on the changes of implied volatility, which can be used as the input features of the LSTM model training. The prediction loss of the model is the smallest when the learning rate is 0.001 and the time step is 140.

3.3 Minimum Variance Delta Hedging Analysis under Different Volatility Models

The minimum variance Delta strategy calculates the minimum variance Delta based on the implied volatility predicted by the quadratic benchmark model of Hull and White (2017) and the LSTM neural network model, respectively, which is completed by Python 3.8.8. pandas and numpy libraries are used for data processing and calculation, and statsmodels library is used for quadratic model establishment.

3.3.1 Hull and White Minimum Variance Delta Hedging Analysis

Since there are more than one option variety corresponding to CSI 300ETF on the same day, the options are first classified according to the code and then the ETF and option price change rates are calculated respectively. The dropna() function is used to eliminate all missing rows in the data, and then the parameter regression of the quadratic benchmark equation can be carried out. The processed data are shown in Table 6:

Table 6. Data Display of Hull & White Quadratic Model

Trading day	Option code	f	Delta	Vega	S	T	△S	△f
2019-12-24	10002117	0.406	0.922	0.002	3.998	0.079	0.005	0.023
2019-12-25	10002117	0.403	1.000	0.000	4.003	0.077	0.001	-0.003
2019-12-26	10002117	0.439	0.931	0.002	4.035	0.074	0.008	0.036
……								
2023-12-29	10006535	0.212	-0.784	0.004	3.447	0.167	0.002	-0.010
2023-12-29	10006536	0.2976	-0.883	0.003	3.447	0.167	0.002	-0.012

Note: Due to the limitation of the table, only the results after three decimal places are reserved, the actual number is six

Then, the training set and the test set of the new data are still divided according to the proportion of 80% and 20%. The parameters a, b and c of the model are 0.0003, -1.0002 and 0.1401 respectively, and then the standard errors under the Black-Scholes Delta and the quadratic benchmark model are calculated respectively. The standard errors of the quadratic benchmark model need to use the regression equation containing parameter values, and the specific calculation formula is as follows:

$$\varepsilon_{HW} = \Delta f - \delta_{bs}\Delta S - \frac{v_{bs}\Delta S}{S\sqrt{T}}(0.0003 - 1.0002\delta_{bs} + 0.1401\delta_{bs}^2) \quad (13)$$

After calculating the error terms of the two respectively, the sum of the squares of the residual error is calculated for the Gain index calculation, and the Gain is -0.0006, when the Gain is less than 0, It shows that in the 977 trading days from December 23, 2019 to December 31, 2023, the return of using Black-Scholes Delta hedging on the whole market of CSI 300ETF options is slightly better than that of Hull and White's minimum variance Delta of about 0.06%. The results are similar to those of Wang Qingshui (2019) on the Taiwan index option market in China. However, since the difference in Gain only occurs when four decimal places are reserved, and it is about 0 before three decimal places are reserved, it can be considered that the portfolio returns under the two hedging strategies are equivalent.

3.3.2 Analysis of Minimum Variance Delta Hedging Under LSTM Neural Network Model

Since the CSI 300ETF has more than one option variety that can be used for hedging on the same day, the minimum variance deltas of all CSI 300ETFs on the same day are considered to be summed and averaged in the process of calculation and mapping, and the minimum variance deltas of each trading day from December 23, 2019 to December 31, 2023 are finally obtained.

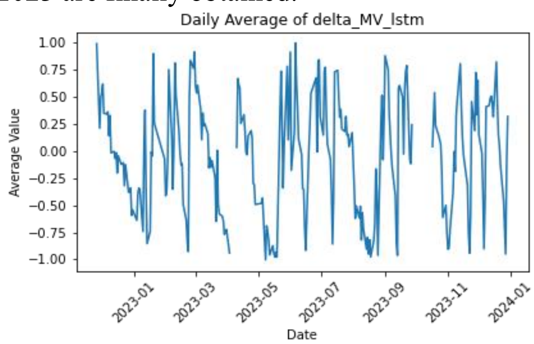

Figure 7. Daily Average Minimum Variance Delta Under the LSTM Model

In Figure 7 and Figure 8, it can be seen that the daily average minimum variance Delta under the LSTM model has the same basic trend as the Black-Scholes Delta, but the variation range of Black-Scholes Delta is -0.3 to 0.1, while the variation range of LSTM minimum variance Delta is -1 to 1. It shows that the variation range of Delta in the same trading day is larger under the LSTM model, and it is more sensitive to the changes of underlying price and expected volatility and more accurate to describe. In the face of the extreme values and inflection points of the Black-Scholes Delta, the minimum variance Delta under the LSTM model shows more drastic fluctuations, which indicates that the variation of Delta after considering the change of implied volatility is significantly larger than that only considering the change of the underlying price, and this factor should be actively taken into account in hedging practice.

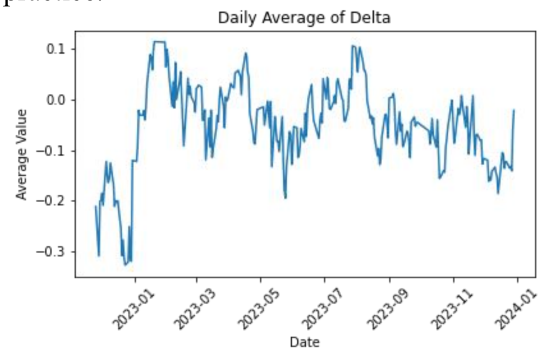

Figure 8. Daily Average Delta Under the Black-Scholes Model

The error terms of the minimum variance Delta under the LSTM model and the minimum variance Delta under the Black-Scholes model are calculated respectively, then the sum of the squares of the residual is calculated for the Gain index calculation, and the Gain is 0.0012, when the Gain is greater than 0. This shows that in the 977 trading days from December 23, 2019 to December 31, 2023, the return of minimum variance Delta hedging after LSTM model is used to forecast implied volatility in the whole option market of Shanghai and Shenzhen 300ETF is slightly better than that of Black-Scholes Delta by about 0.12%. There is a significant difference in Gain when decimal places are reserved, and it can be considered that the portfolio return of the minimum variance Delta hedging strategy of LSTM model is significantly better and it can be intuitively seen that the minimum variance Delta hedging effect under the LSTM neural network model is the best. The second is the Black-Scholes Delta hedge, and the least effective is the minimum variance Delta under the Hull and White quadratic model.

4. Conclusion

This study constructs minimum-variance Delta hedging models under different volatility forecasting frameworks and conducts empirical analysis using CSI 300 ETF option trading data. The main conclusions are as follows:

1. The implied volatility curve of CSI 300 ETF options exhibits a pronounced smile

pattern. Among the tested models, the LSTM demonstrates superior capability in capturing volatility time-series characteristics compared to the CNN, which shows overfitting phenomena in sequential data. Empirical results indicate that eight input features—including option closing price, strike price, and time to maturity—exhibit significant explanatory power for implied volatility. The LSTM model achieves optimal forecasting performance under a learning rate of 0.001 and a time step of 140.

2. Comparative analysis based on the hedging effectiveness metric Gain reveals that the minimum-variance Delta hedging strategy derived from the LSTM volatility forecasting model outperforms both the Black-Scholes Delta and Hull & White minimum-variance Delta strategies, with risk management effectiveness improved by 0.12% and 0.18%, respectively. This suggests that integrating LSTM-based volatility predictions into Delta calculations significantly enhances hedging efficiency.

3. Compared to the Black-Scholes Delta, the minimum-variance Delta generated by the LSTM model exhibits higher sensitivity to changes in underlying prices and expected volatility, particularly at extreme points and inflection points. This characteristic indicates that the minimum-variance Delta framework, which incorporates implied volatility forecasting, is more suitable for practical option hedging scenarios.

Acknowledgments: This paper is supported by the 2023 key project of the "14th Five-Year Plan" of Education Science of Heilongjiang Province. Project number: GJB1423125

Reference

[1] Cao J, Chen J, Farghadani S, et al. Gamma and Vega hedging using deep distributional reinforcement learning. Frontiers in Artificial Intelligence, 2023, 6: 1129370.

[2] Chen Y C, Huang W C. Constructing a stock-price forecast CNN model with gold and crude oil indicators. Applied Soft Computing, 2021, 112: 107760.

[3] Crépey S. Delta-hedging vega risk? Quantitative Finance, 2004, 4(5): 559-579.

[4] Kirisci M, Cagcag Yolcu O. A new CNN-based model for financial time series: TAIEX and FTSE stocks forecasting. Neural Processing Letters, 2022, 54(4): 3357-3374.

[5] Hailiang Zhang, Muhammad Atif Sattar, Haijun Wang. Uncertainty measure: As a proxy for the degree of market imperfection. International Review of Economics & Finance.2024,89:159-171

[6] Yadav A, Jha C K, Sharan A. Optimizing LSTM for time series prediction in Indian stock market. Procedia Computer Science, 2020, 167: 2091-2100.

[7] Yang M, Wang J. Adaptability of financial time series prediction based on BiLSTM. Procedia Computer Science, 2022, 199: 18-25.

[8] Yong'an Zhang, Binbin Yan, Memon Aasma. A novel deep learning framework: Prediction and analysis of financial time series using CEEMD and LSTM. Expert systems with applications, 2020, 159: 113609.

[9] Kun Xia, Xuewei Yanga, Peng Zhu. Delta hedging and volatility-price elasticity: A two-step approach. Journal of Banking and Finance. 2023, 153.106898

[10] Li X, Zhou C, Huang W. Delta-hedged gains of SSE 50 ETF options. Applied Economics

[11] Letters, 2022, 29(20): 1864-1867.

[12] Lina Song, Wang Yu, Yousheng Tan, Ke Duan. Calculations of fractional derivative option pricing models based on neural network. Journal of Computational and Applied Mathematics. 2024,437 .115462

[13] Udayan Sharma, Madhusudan Karmakar. Measuring minimum variance hedging effectiveness: Traditional vs. sophisticated models. International Review of Financial Analysis. 2023, 87 .102621

[14] Alexander C, Rubinov A, Kalepky M, et al. Regime-dependent smile-adjusted delta hedging. Journal of Futures Markets, 2012, 32(3): 203-229.

[15] Hull J, White A. Optimal delta hedging for options. Journal of Banking & Finance, 2017, 82: 180-190.

[16] Xiaoyijing Chen, Siyuan Liu, Zailin Xu, Mei Yu. Asymmetry in option implied

volatility and yield: Evidence from China's ETF options market Pacific-Basin Finance Journal.2024,85. 102386

[17] Yu-Sheng Lai. Economic evaluation of dynamic hedging strategies using high-frequency data. Finance Research Letters. 2023, 57. 104230

