上海交通大学学位论文

中国市场的期货期权的校准

姓　　名：任璨

学　　号：520070910098

导　　师：Samuel Drapeau

学　　院：数学科学学院

专业名称：统计学

申请学位层次：学士

2024 年 02 月

A Dissertation Submitted to

Shanghai Jiao Tong University for Bachelor's Degree

CALIBRATION OF COMMODITY OPTIONS IN CHINA

Author: Ren Can

Supervisor: Samuel Drapeau

School of Mathematical Sciences

Shanghai Jiao Tong University

Shanghai, P.R. China

February, 2024

上海交通大学

学位论文原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全知晓本声明的法律后果由本人承担。

学位论文作者签名：

日期：　　年　　月　　日

上海交通大学

学位论文使用授权书

本人同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。

本学位论文属于：

□**公开论文**

□**内部论文**，保密□1 年/□2 年/□3 年，过保密期后适用本授权书。

□**秘密论文**，保密____ 年（不超过 10 年），过保密期后适用本授权书。

□**机密论文**，保密____ 年（不超过 20 年），过保密期后适用本授权书。

（请在以上方框内选择打“✓”）

学位论文作者签名：　　指导教师签名：

日期：　　年　　月　　日　　日期：　　年　　月　　日

摘 要

本文探讨了我对考虑交易流动性的期货期权定价问题的研究。传统的 Black-Scholes 期权定价模型假设标的物的隐含波动率为常数，然而这一假设并不准确。因此，我引入了随机波动率模型，我讨论了随机波动率模型的发展过程，其中 SABR 校准是一种常用且完善的随机波动率模型。在 SABR 校准的基础上，我提出了交易量加权的 SABR 校准方法，并且结合实际数据分析了这样做的原因。我通过修改后的理论模型计算修正后的 SABR 参数，并使用 ARIMA 和 LSTM 两种时间序列模型对获得的参数的时间序列进行建模进而进行预测，通过预测的参数来对未来期权价格进行预测。最后，我设计了一种考虑交易流动性的交易策略进行回测。结果表明，采用考虑交易流动性的 SABR 校准模型，在实际期货期权交易中能够获得更多的超额收益，表现出明显的优势。

关键词：期权定价，Black-Scholes，SABR，ARIMA，LSTM

ABSTRACT

This paper examines my research on the pricing of futures options considering trading liquidity. The traditional Black-Scholes options pricing model assumes a constant implied volatility for the underlying asset, which is not accurate. Therefore, I introduce a stochastic volatility model and discuss its development process, with the SABR calibration being a commonly used and refined stochastic volatility model. Building upon the SABR calibration, I propose a trading volume-weighted SABR calibration method and analyze the reasons for doing so with actual data. I compute the adjusted SABR parameters using the modified theoretical model and utilize two time-series models, ARIMA and LSTM, to model the time series of obtained parameters for forecasting. I then use the forecasted parameters to predict future option prices. Finally, I design a trading strategy considering trading liquidity for backtesting. The results demonstrate that employing the SABR calibration model considering trading liquidity leads to greater excess returns in actual futures options trading, showcasing a significant advantage.

Key words: option pricing，Black-Scholes，SABR，ARIMA，LSTM

Contents

Chapter 1 INTRODUCTION

1.1 Foreword

In the financial trading market, options, as a type of financial derivative, are also an integral component of the financial market. Depending on the underlying asset, options come in various forms. This paper primarily focuses on futures options, which are options with futures contracts as their underlying asset. Futures options grant the holder the right, but not the obligation, to buy (call option) or sell (put option) a futures contract at a predetermined price in the future. Compared to traditional futures contracts, futures options offer greater flexibility to investors. Investors can choose to exercise their rights before the contract expires, allowing them to react to market conditions and personal preferences. This flexibility makes futures options an essential tool for investors engaged in risk management and speculative trading.

Similar to stocks, profits can also be gained through options trading. However, unlike stock prices, option pricing is more complex, making predicting option prices more challenging. The Black-Scholes formula is a well-known option pricing formula that provides a rational trading price for options based on factors such as the option's strike price, expiration date, underlying asset's forward price, risk-free interest rate, and volatility. With option pricing formulas, we can make reasonable and effective predictions of option prices. In the quantitative trading industry, predicting the price trends of financial derivatives is an essential process, and investors can profit from these predictions by making informed investment decisions based on market conditions and risk preferences.

Thorough research and refinement of option pricing models can lead to more accurate assessment of option prices, enabling investors to seize greater investment opportunities in the financial market. Therefore, in-depth research on option pricing is highly meaningful.

1.2 The main content of this paper

In this study, I focus on the problem of option pricing in the Chinese options trading market, with due consideration to trading liquidity, with a specific focus on futures options.

My primary approach involves leveraging daily cross-sectional data of futures options to obtain trading volumes and prices for both call and put futures options at various strike prices. Utilizing the Black-Scholes formula, I am able to calculate option prices based on the strike price of futures options, the forward price of the underlying futures at the exercise date, the risk-free interest rate, and volatility, where volatility is the only metric not directly obtainable. Naturally, this leads to the idea of inferring volatility by examining past trading data of futures options and their actual transaction prices, employing the Black-Scholes model. Subsequently, by constructing a volatility prediction model, I aim to forecast futures options prices.

My specific research process is as follows. By examining the statistical relationship between daily trading volumes of futures options and their underlying futures, I aimed to identify dates characterized by significant trading volumes for both instruments.Following this, on these selected dates, I compute implied volatilities of futures options with different strike prices using the Black-Scholes model, thereby elucidating market volatility dynamics. Post-analysis, I conduct meticulous SABR calibration, considering subtle differences in the trading volume weight ratio, to enhance the futures option pricing model. Additionally, I utilize the calibrated parameters derived from SABR calibration to train an LSTM model. This LSTM (Long Short-Term Memory) model is employed to forecast SABR parameters for the following day, facilitating the computation of actual futures option prices across a range of strike prices based on the predicted SABR parameters.

1.3 The significance of this article

In comparison to previous studies on option pricing, my research takes into consideration the liquidity of futures options to calibrate volatility. This undoubtedly leads to more accurate and practically meaningful results. In real-life scenarios, it is preferable to choose options with higher liquidity for quantitative trading. By employing deep learning models to analyze historical data, I can develop a more precise model capable of forecasting future volatility more accurately. My research on option pricing also aids traders in making more accurate assessments of market conditions, thus holding significant implications.

Specifically, in the option pricing process, when calculating the volatility of options for each day, we observe a smile curve in volatility across different strike prices, known as the

volatility smile. This phenomenon arises because options deep in or out of the money exhibit more significant volatility, contrary to the flat curve predicted by the Black-Scholes model. Consequently, the SABR model introduces more complex parameters on top of the Black-Scholes model to fit the volatility smile. However, in the Chinese options trading market, trading volumes for options at most strike prices are relatively low. Hence, I suggest focusing the calibration of the SABR model on options with higher trading volumes to obtain more accurate results on options with better liquidity.

1.4 Literature Review

1.4.1 The Black-Scholes Option Pricing Model

Options are financial derivatives that grant the holder the right, but not the obligation, to buy (call option) or sell (put option) a specified quantity of an underlying asset at a predetermined price at a specific time or date in the future. Due to their unique nature, options have been the subject of extensive research in financial mathematics.

The Black-Scholes model[1-2], introduced by Fischer Black and Myron Scholes in 1973, is a landmark in the pricing of European options. It provides a mathematical framework to calculate the theoretical price of options based on a set of assumptions, including the stock price following a geometric Brownian motion and the market operating without frictions.

While the Black-Scholes model is widely used for stock options, Fischer Black extended the model to address the pricing of futures options in 1976, known as the Black-76 model[3]. Futures options have unique characteristics that differentiate them from stock options, such as the obligation to trade the underlying asset at a future date at a set price and the requirement of a margin.

Implied volatility is the volatility figure that is inferred by substituting the market trading price of an option into a theoretical option pricing model, such as the Black-Scholes (BS) model. Theoretically, deriving the implied volatility of an option is not a complex task. Since option pricing models, like the BS model, provide a quantitative relationship between the option price and five fundamental parameters (underlying asset price, strike price, interest rate, maturity, and volatility), by inputting the first four parameters along with the market trading price of the option into the pricing formula, the only unknown quantity—implied

volatility—can be resolved.

Implied volatility reflects the valuation level of an option. There is a significant difference in prices among options with different strike prices and maturities, making direct comparison impossible. It is only through the pricing formula that the market price of an option is deduced to obtain the implied volatility, which can represent the true value of the option.

The Black-Scholes model assumes that the volatility of the underlying asset is a constant, suggesting that options of the same underlying asset should have the same implied volatility. However, empirical studies have shown that for options with the same underlying asset and maturity date, the more the strike price deviates from the spot price, the higher the implied volatility tends to be[4]. This pattern of implied volatility across different strike prices resembles a smile, hence the term ”volatility smile.”

Hull, in his book, detailed how implied volatility should be modeled and calculated, offering profound insights into market practices[5]. Rubinstein, through his research, conducted empirical analysis on the volatility smile phenomenon, demonstrating the variation in options' implied volatility across different strike prices[4]. Furthermore, Yan's study further explored the relationship between the slope of the implied volatility smile and future stock returns[6]. Gatheral, in his work, delved into the construction of the volatility surface and its impact on options pricing[7]. Lastly, Dumas and colleagues analyzed the shape of the implied volatility function through empirical testing, providing an important perspective for understanding the information implied by the market[8].

1.4.2 Stochastic Volatility Models

Stochastic Volatility (SV) models have deep roots in both financial econometrics and mathematical finance, with their origins dating back to pioneering work by Clark[9]. Clark's approach modeled asset returns as a stochastic process of information, which led to time-varying models that approximate the distribution of asset returns. This foundational work was later expanded upon by Tauchen and Pitts[10], who proposed a mixed distribution model for returns, taking into account the timing of information arrival.

The financial econometrics field was significantly advanced with the introduction of the AutoRegressive Conditional Heteroskedasticity (ARCH) model by Engle in 1982[11]. The ARCH model was instrumental in addressing the challenge of heteroskedasticity in financial

time series data by characterizing the conditional variance at time t as a linear combination of past squared errors. This innovation allowed for the estimation of future volatility based on historical data and was pivotal in capturing the volatility clustering phenomenon observed in financial markets.

Despite its contributions, the ARCH model faced limitations, including the need for non-negative parameters to ensure stability and increasing computational complexity with more lags. To address these issues, Bollerslev[12] introduced the Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) model in 1986. The GARCH model advanced volatility forecasting by incorporating past squared residuals and past conditional variances, offering a more nuanced and computationally efficient method for volatility modeling.

The GARCH model's success in balancing model complexity with forecasting accuracy has established it as a cornerstone in financial volatility research. It simplifies the estimation process compared to high-order ARCH models and provides a robust mechanism for predicting volatility that considers the dynamic nature of financial markets.

Building on the earlier frameworks, Hull and White[13] focused on European option pricing, assuming a continuous-time stochastic volatility model and introducing a diffusion expression for volatility. They suggested that the volatility of the underlying asset follows a positive diffusion process. In parallel, Taylor[14] proposed a non-continuous time volatility model as an alternative to the ARCH model, providing a fresh perspective on the dynamics of volatility.

The integration of SV models into options pricing discussions was furthered by the contributions of Scott[15], Hull and White[13], and Wiggins[16]. These scholars highlighted the potential of SV models to capture complex financial market characteristics, such as volatility clustering and the leptokurtic nature of asset returns.

The SABR model, a stochastic volatility model, was proposed by Hagan et al. in 2002[17] as a valuable complement to the constant volatility assumption in the Black-Scholes model. With its relaxed assumptions, simple structure, and the availability of an analytical solution, the SABR model is capable of adapting to various market environments and provides a good depiction of the shape of the implied volatility curve. It serves as an effective model for managing the "volatility smile" and has been extensively studied and used in the international industry since its introduction by Hagan et al. in 2002.

1.5 Summary

In this study, I investigate option pricing in the Chinese options trading market, focusing on futures options. By analyzing daily cross-sectional data, I employ the Black-Scholes formula to calculate option prices, with a particular emphasis on inferring volatility. Through statistical analysis and meticulous SABR calibration, I refine the pricing model, leveraging advanced techniques such as LSTM for forecasting. Ultimately, this research aims to provide more accurate and practical insights into option pricing, enhancing decision-making for investors and traders in the financial market.

Chapter 2 Volatility Smile in Options Pricing

2.1 Black-Scholes Model and Futures Options Pricing

2.1.1 Black-Scholes Model

The Black-Scholes option pricing model[1-2] is based on several fundamental assumptions:

1. The spot price S_t follows a lognormal process with constant volatility σ.
2. Short selling is permissible.
3. There are no transaction costs or taxes.
4. The risk-free rate r is constant across all maturities.
5. There is no riskless arbitrage.
6. Trading is continuous between the present time ($t = 0$) and the expiry time ($t = T$).

It is assumed that the stock price follows a lognormal distribution, represented by the stochastic differential equation:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \quad ([1]) \tag{2–1}$$

Here, W_t represents the standard Wiener process.

For a European call option C, it is a function of the underlying stock price S_t and time t, i.e., $C = C(S_t, t)$. Applying Itô's lemma, I obtain:

$$dC = \left(\frac{\partial C}{\partial S_t}\mu S_t + \frac{\partial C}{\partial t} + \frac{1}{2}\frac{\partial^2 C}{\partial S_t^2}\sigma^2 S_t^2\right)dt + \frac{\partial C}{\partial S_t}\sigma S_t dW_t \tag{2–2}$$

Constructing a Delta hedge portfolio P (short 1 unit of the option and long $\frac{\partial C}{\partial S_t}$ units of the stock) yields:

$$dP = \left(-\frac{\partial C}{\partial t} - \frac{1}{2}\frac{\partial^2 C}{\partial S_t^2}\sigma^2 S_t^2\right)dt \tag{2–3}$$

In a market without riskless arbitrage, the portfolio's rate of return over dt must equal the risk-free rate r, i.e., $dP = rPdt$. This leads to the Black-Scholes partial differential equation:

$$\frac{\partial C}{\partial t} + rS_t\frac{\partial C}{\partial S_t} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 C}{\partial S_t^2} = rC \tag{2–4}$$

According to risk-neutral pricing theory, the expected return rate of the underlying stock S_t is the risk-free rate r, and the discount rate for the option is also r. Thus, the option price

C at the current time, assuming the exercise time is T and the current time is $t = 0$, with strike price K, is given by:

$$C = e^{-rT}\mathbb{E}[\max(S(T) - K, 0)] \tag{2–5}$$

The prices of European call and put options are calculated as follows, where $N(.)$ is the cumulative distribution function of the standard normal distribution:

$$C = S_0 N(d_1) - Ke^{-rT}N(d_2) \tag{2–6}$$

$$P = Ke^{-rT}N(-d_2) - S_0 N(-d_1) \tag{2–7}$$

where

$$d_1 = \frac{\ln(S_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}} \tag{2–8}$$

$$d_2 = \frac{\ln(S_0/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T} \tag{2–9}$$

2.1.2 Pricing of futures options

The primary focus of this study is on futures options, which are options with futures contracts as the underlying assets. Unlike stock options, futures exhibit several distinctions. Firstly, futures contracts obligate the buyer to purchase or sell a specific asset at an agreed price on a specified future date, with corresponding obligations for the seller. Therefore, the price of a futures contract reflects the future price of the underlying asset, rather than its current price. Secondly, constructing futures contracts typically requires only a small portion of the contract price to be provided as margin. These disparities lead to differences in the pricing of futures options compared to stock options. Fischer Black introduced the pricing formula for futures options in 1976, also known as the Black-76 model[3].

The pricing process of futures options differs from that of stock options. In the construction of the Black-Scholes differential equation for futures options, the cost of constructing futures contracts is assumed to be zero ($F_t^c = 0$ represents the cost of constructing the futures contract at time t), implying $\frac{\partial C}{\partial F_t^c} F_t^c = 0$. This assumption is based on the absence of arbitrage opportunities. Therefore, according to the assumption of no riskless arbitrage, I have:

$$dP = -rCdt \tag{2–10}$$

When time $t = T$ (the expiration time) arrives, the option price C must satisfy $C = \max(F(T) - K, 0)$, where K is the strike price. The solution to the partial differential equations yields (F_t represents the forward price of the futures contract at time t):

$$C = Fe^{-rT}N(d_1) - Ke^{-rT}N(d_2) \tag{2–11}$$

$$P = Ke^{-rT}[1 - N(d_2)] - Fe^{-rT}[1 - N(d_1)] \tag{2–12}$$

where

$$d_1 = \frac{\ln(F/K) + \sigma^2 T/2}{\sigma\sqrt{T}} \tag{2–13}$$

$$d_2 = \frac{\ln(F/K) - \sigma^2 T/2}{\sigma\sqrt{T}} \tag{2–14}$$

Based on the above, I can derive the pricing equation for futures options as follows:

$$price = f(F_t, K, r, t, T, \sigma, flag) \tag{2–15}$$

2.2 Volatility Smile

2.2.1 Implied Volatility

For traded options, I can obtain their trading price ($price$), option type ($flag$), strike price (K), underlying futures price (F), expiration date (T), and risk-free interest rate (r). Thus, apart from volatility (σ), all other parameters can be acquired. Clearly, through our option pricing derivation process, for each specific option type, strike price, underlying futures price, expiration date, risk-free interest rate, and volatility, I are able to uniquely compute the option's price.

Firstly, there exists an equality relationship between call options and put options with the same strike price. According to the principle of risk-neutral pricing, I have $C = e^{-rT}\mathbb{E}[\max(F(T) - K, 0)]$ and $P = e^{-rT}\mathbb{E}[\max(K - F(T), 0)] = -e^{-rT}\mathbb{E}[\min(F(T) - K, 0)]$. So, $C - P = e^{-rT}\mathbb{E}[F(T) - K] = e^{-rT}(F - K)$. Then, this is the Put-call parity; that is to say, as long as I prove that the option price is strictly monotonic with respect to σ for call options, the price of put options naturally also varies strictly monotonically with σ.

Prior to this, I assume: $F > 0$, $K > 0$, $\sigma > 0$, $T > 0$. **i**: For $F = K$, i.e., at-the-money option, I can observe that $C = P = F\mathrm{e}^{-rT}(\mathrm{N}(d_1) - \mathrm{N}(d_2))$ where $d_1 = -d_2 = \frac{\sigma\sqrt{T}}{2}$. It is

easy to see that as σ increases, d_1 increases while d_2 decreases. Therefore, the prices of both call and put options strictly increase. **ii**: When F is not equal to K, the price of the option is a function of σ when all parameters except σ are fixed. Moreover, it is evident that this function has a derivative (Here f is the pdf of the standard normal distribution).

$$\begin{aligned}\frac{\partial C}{\partial \sigma} &= Fe^{-rT} f\left(d_1\right)\left(\frac{\sqrt{T}}{2} - \frac{\ln(F/K)}{\sigma^2\sqrt{T}}\right) - Ke^{-rT} f\left(d_2\right)\left(-\frac{\sqrt{T}}{2} - \frac{\ln(F/K)}{\sigma^2\sqrt{T}}\right) \\ &= Fe^{-rT} f\left(d_1\right)\frac{\sqrt{T}}{2} + Ke^{-rT} f\left(d_2\right)\frac{\sqrt{T}}{2} + e^{-rT}\frac{\ln(F/K)}{\sigma^2\sqrt{T}}\left(Kf\left(d_2\right) - Ff\left(d_1\right)\right)\end{aligned}$$

If I can prove that $\ln(F/K)(Kf(d_2) - Ff(d_1)) \geq 0$, it would indicate that, given other parameters, the price of the call option C strictly increases with σ. I can further simplify $Kf(d_2) - Ff(d_1)$ as follows:

$$\begin{aligned}Kf(d_2) - Ff(d_1) &= \frac{1}{\sqrt{2\pi}}\left(Ke^{-\frac{(\ln(F/K)-\sigma^2T/2)^2}{2\sigma^2T}} - Fe^{-\frac{(\ln(F/K)+\sigma^2T/2)^2}{2\sigma^2T}}\right) \\ &= \frac{K}{\sqrt{2\pi}}\left(e^{-\frac{(\ln(F/K)-\sigma^2T/2)^2}{2\sigma^2T}} - \frac{F}{K}e^{-\frac{(\ln(F/K)+\sigma^2T/2)^2}{2\sigma^2T}}\right)\end{aligned}$$

Then, I can denote F/K as m ($m > 0$), and obtain

$$\begin{aligned}Kf(d_2) - Ff(d_1) &= \frac{K}{\sqrt{2\pi}}\left(e^{-\frac{(\ln(m)-\sigma^2T/2)^2}{2\sigma^2T}} - me^{-\frac{(\ln(m)+\sigma^2T/2)^2}{2\sigma^2T}}\right) \\ &= \frac{K}{\sqrt{2\pi}}\left(e^{-\frac{(\ln(m)-\sigma^2T/2)^2}{2\sigma^2T}} - e^{-\frac{(\ln(m)+\sigma^2T/2)^2}{2\sigma^2T}+\ln(m)}\right)\end{aligned}$$

Next, I need to compare $-\frac{(\ln(m)-\sigma^2T/2)^2}{2\sigma^2T}$ and $-\frac{(\ln(m)+\sigma^2T/2)^2}{2\sigma^2T} + \ln(m)$.

$$-\frac{(\ln(m) - \sigma^2T/2)^2}{2\sigma^2T} + \frac{(\ln(m) + \sigma^2T/2)^2}{2\sigma^2T} - \ln(m)$$

Taking the difference, I have

$$-\frac{(\ln(m) - \sigma^2T/2)^2}{2\sigma^2T} + \frac{(\ln(m) + \sigma^2T/2)^2}{2\sigma^2T} - \ln(m) = 0$$

Thus, $\ln(F/K)(Kf(d_2) - Ff(d_1)) = 0$, which leads to

$$\frac{\partial C}{\partial \sigma} = Fe^{-rT} f(d_1) \frac{\sqrt{T}}{2} + Ke^{-rT} f(d_2) \frac{\sqrt{T}}{2}$$

Therefore,

$$\frac{\partial C}{\partial \sigma} > 0$$

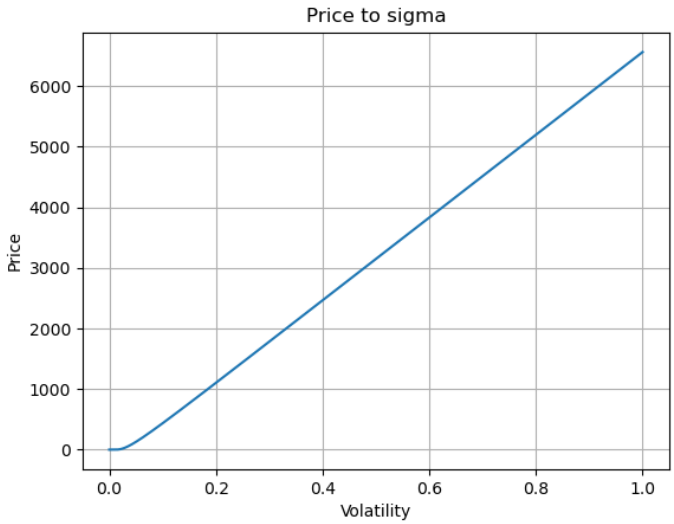


Figure 2–1 Price to sigma

I plotted the call option price as a function of volatility using the Black-Scholes pricing formula based on the data of CU2403 futures options on February 2, 2024. From Figure 2–1, it is evident that the option price increases monotonically with increasing volatility. As the call option price C strictly increases with volatility given all other parameters, and varies continuously with σ, for a given option price and all other parameters, I can uniquely determine a value for volatility; this is referred to as implied volatility.

2.2.2 Time-Varying Implied Volatility

The Black-Scholes model posits that the underlying asset's price dynamics are governed by a log-normal distribution, with the volatility parameter denoting the constant volatility inherent to the log-normal process.

Theoretically, for a given option, the implied volatility (IV) derived from different trading dates should be a constant value. I have plotted the implied volatilities calculated on various

Figure 2–2 Price to date

trading days for the CU2403 option with a strike price of 70,000. Upon examination of the figure 2–2, it is evident that the implied volatility does not maintain a constant value; this inconsistency indicates that the assumption of constant implied volatility within the Black-Scholes model is not entirely accurate.

2.2.3 Volatility smile

Based on the Assumptions of the Black-Scholes Formula, the implied volatility calculated on the same trading day should exhibit a flat line as the strike price varies.However, if I compute the implied volatility for different strike prices using actual data, for example, using the data of CU2403 futures options on February 2, 2024, and employing the Newton-Raphson method to solve for the implied volatility based on the Black-Scholes pricing formula, I obtain the implied volatility versus strike price plot as shown in the figure 2–3. It can be observed that in reality, the implied volatility of the option is not constant for different strike prices. This phenomenon is commonly referred to as the volatility smile.

In the context of the volatility smile phenomenon, I observe that for options deviating from the forward price (i.e., options deep in-the-money and deep out-of-the-money), their implied volatility is higher than the constant value assumed by the Black-Scholes model. It has been demonstrated that option prices are monotonic increasing functions of asset volatil-

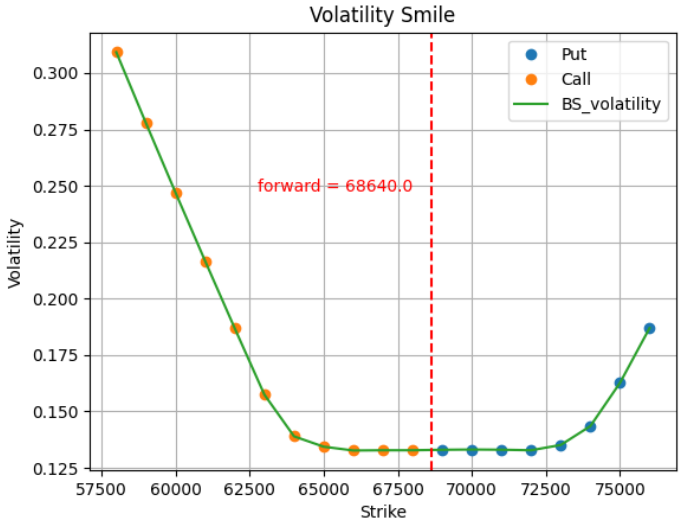


Figure 2–3 Volatility Smile

ity, implying that option prices are significantly overestimated in the deep in-the-money and deep out-of-the-money regions.

Why does this phenomenon occur? In fact, before the stock market crash of 1987, implied volatility of options did indeed form a flat line. However, after the crash, the volatility smile emerged. This suggests that post-crash, traders realized that the market exhibits significant skewness and that dramatic events are possible. Investors tend to overestimate the probability of rare events, leading to an overestimation of the expected value of options and ultimately resulting in the volatility smile phenomenon.

2.3 Option Liquidity

2.3.1 Main Contract

For each day, there are multiple futures contracts with the same underlying commodity. For example, CU2401 and CU2403 are both futures contracts based on the commodity CU, differing only in their expiration dates. Consequently, for CU futures options, there are multiple option contracts available each day, with varying trading volumes. I refer to the contract with the highest trading volume on a given day as the ”main contract”. As shown in the figure 2–4, empirical data reveals that the main contract typically accounts for nearly 50% of

the total trading volume of all futures contracts.

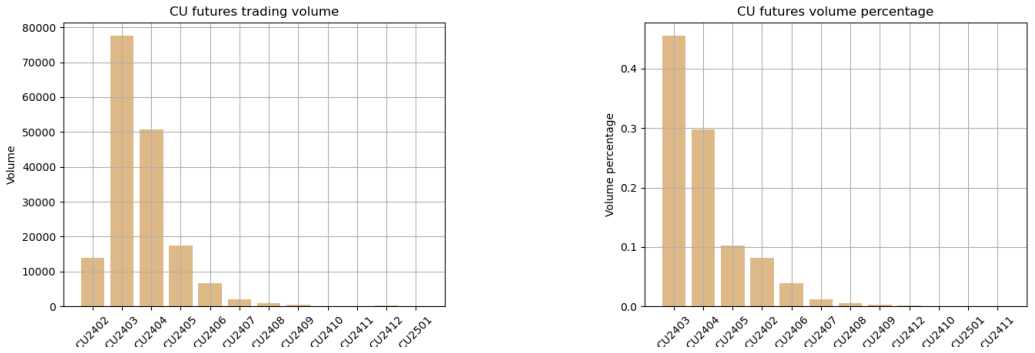
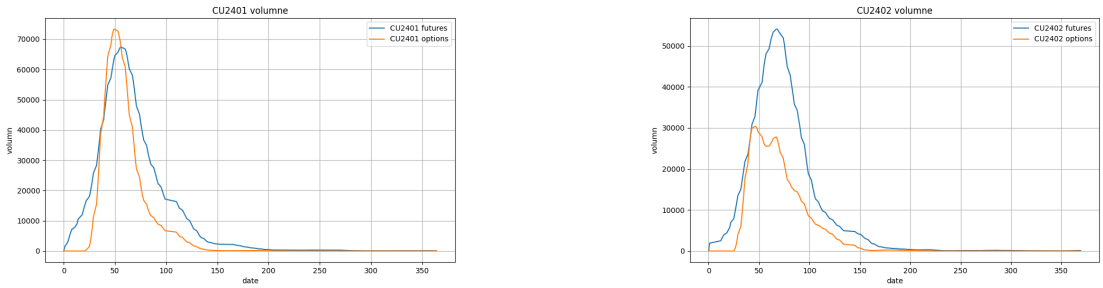
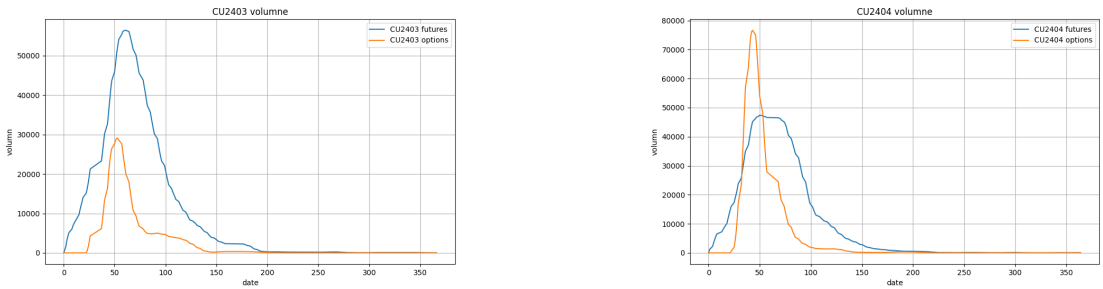


Figure 2–4 Futures trading volume



(a) CU2401 (b) CU2402



(c) CU2403 (d) CU2404

Figure 2–5 CU futures vs option volume(smoothed)

The central premise of this article is that only trading data with sufficient volume contains meaningful information.Furthermore, there is a strong linear correlation between the trading volume of futures options and the corresponding trading volume of futures contracts. The research investigated the correlation coefficient between the trading volumes of CU futures

options and their corresponding underlying assets from 2021 to April 2024. The average correlation coefficient was found to be 0.81, indicating a strong correlation between them. As shown in the figure 2–5, it can be observed that when the futures trading volume is high, the options trading volume is also high. Therefore, I will select the trading data corresponding to the futures options of the main and second contracts each day for modeling purposes.

2.3.2 Liquidity Among Different Strike Prices

For each day, the trading prices and volumes vary across different strike prices of the same futures option. As illustrated in the graph 2–6, the trading volumes for deep in-the-money and deep out-of-the-money futures options are close to zero. Considering the central theme of this paper that only information with trading volume is valuable, I will first eliminate option trading data with trading volumes accounting for less than 5% in both call and put options. These trading data with unreasonable prices will lead to incorrect conclusions. Instead, I will select option contracts with trading volumes accounting for more than 5% in both call and put options as effective contracts to calculate implied volatility and make predictions.

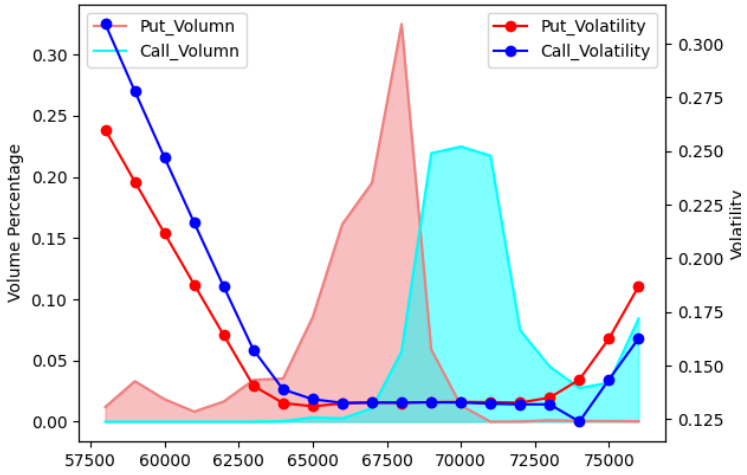


Figure 2–6 Volatility vs Volume

For futures options with a trading volume share greater than 3%, I inferred their implied volatility using the Black-Scholes formula as depicted in the figure 2–7. After removing the insufficient trading volume data, I observed that 95.5% of the remaining options represent

call options and 92.8% represent put options, indicating that the filtered data constitutes the majority of the options traded on the given dates. Furthermore, the remaining options no longer include those deeply out-of-the-money options, resulting in a narrower range of implied volatility changes. Most of the implied volatilities corresponding to the exercise prices now vary within a very small range.

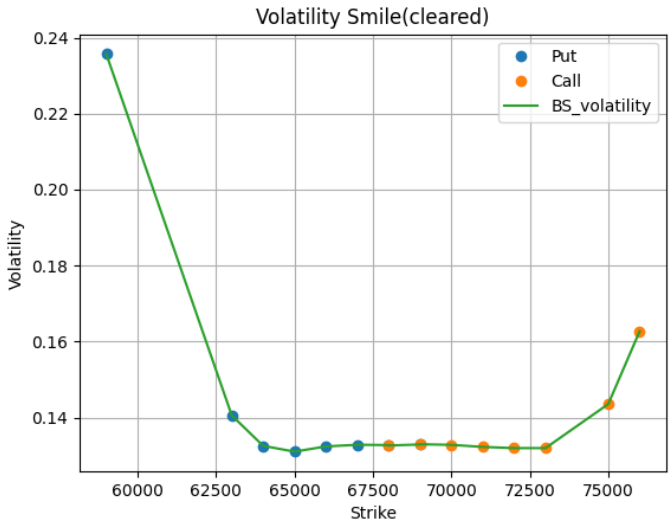


Figure 2–7 Volatility Smile(cleared)

2.3.3 Liquidity Issues

Here still exists a question: why is the liquidity of futures options trading worth considering? Option liquidity refers to the ability of an asset to be bought and sold in the market at reasonable prices. My understanding is that I can predict future option prices based on my model and method, but I can never know the true value of future option prices accurately. This is because there is a noise term involved, which will inevitably cause actual prices to fluctuate around my predicted prices. What I can do is to predict the part outside the noise term as accurately as possible. In actual options trading, options with low liquidity, such as deep in-the-money and deep out-of-the-money options, have very poor liquidity, which means that option trading prices may deviate significantly from my predicted values, indicating that the proportion of the noise term will be large. When formulating trading strategies, it is difficult to control the prices of these options with low liquidity. Therefore, I believe it is

better to put more weight on options with good liquidity for SABR calibration. This way, our calibration model will be more stable and I can achieve more accurate predictions in options with high liquidity.Reflected in the trading strategy, I will assign greater position weights to futures options at strike prices with better liquidity, while holding fewer or no positions in options with poor liquidity.

Chapter 3 SABR Calibration

3.1 Stochastic volatility model

Due to the Black-Scholes model's difficulty in explaining the volatility smile curve, its assumption of constant volatility is often considered inaccurate. A corrective approach involves stochastic volatility models, which assume that implied volatility also follows a random process. In stochastic volatility models, there are two important representatives: the stochastic volatility model proposed by Steve Heston[18] and the SABR model[17], which is another highly successful stochastic volatility forward price model.

3.1.1 Heston Model

The Heston model was proposed by Steven L. Heston in a paper published in 1993. In stochastic volatility models, volatility is not a fixed variable but changes over time. However, this stochastic volatility model alone cannot capture the skewness effect caused by market correlation. The Heston model allows for correlation between asset price movements and volatility, assuming that volatility follows a mean-reverting square root process, i.e., the Cox-Ingersoll-Ross (CIR) process. Heston (1993) applied this model to both the foreign exchange options market and the bond options market. The model can be represented as follows:

$$\begin{aligned} dS(t) &= \mu S dt + \sigma_t S dW_1(t) \\ d\sigma_t &= -\beta \sigma_t dt + \delta dW_2(t) \\ d\left\langle W_1, W_2 \right\rangle &= \rho(t) dt \end{aligned}$$

where $S(t)$ is the asset price, σ_t is the volatility, σ_t^2 follows a CIR process, μ is the risk-free interest rate. According to Itô's lemma,

$$d\sigma_t^2 = \left[\delta^2 - 2\beta\sigma_t^2\right] dt + 2\delta\sigma_t dW_2(t) \tag{3–1}$$

This equation can be written as a square root process (CIR process):

$$d\sigma_t^2 = \kappa(t)\left(\theta(t) - \sigma_t^2\right) dt + \xi(t)\sigma_t dW_2(t) \tag{3–2}$$

where the parameters $\kappa(t), \xi(t), \theta(t)$ are constants, $\kappa(t)$ represents the mean reversion rate, $\xi(t)$ represents the volatility of volatility, $\theta(t)$ represents the long-term mean-reverting

level of σ_t^2, W_1 and W_2 are standard Brownian motions, and $\rho(t)$ represents the correlation coefficient between dW_1 and dW_2.

3.1.2 SABR Model

The acronym SABR stands for Stochastic Alpha Beta Rho. It was introduced by Hagan, Kumar, and Lesniewski in 2002 as they investigated the dynamic changes in forward prices. The specific modeling of the SABR model is as follows:

$$\begin{aligned} dF_t &= \alpha_t F_t^{\beta} dW_t^s \\ d\alpha_t &= v\alpha_t dW_t^{\sigma} \\ \left\langle W_t^s, W_t^{\sigma} \right\rangle &= \rho dt \end{aligned}$$

where, F_t represents the forward price of the underlying asset, α_t denotes the stochastic process component of the underlying asset, W_t^{σ} and W_t^s are standard Brownian motions, and ρ is the correlation coefficient between W_t^s and W_t^{σ}. v represents the volatility of volatility, taking non-negative values.

The SABR model is determined by four parameters, namely α (initial volatility level), β (skewness control), ρ (correlation between asset price and volatility), and ν (volatility of volatility). The parameter β determines the shape of the volatility distribution of the asset price, controlling the relationship between the dynamic behavior of the asset price and volatility. The range of values for β is typically $[0, 1]$, and different values affect the shape of the volatility smile. When $\beta = 1$, the asset price follows a lognormal distribution; when $\beta = 0.5$, the volatility follows the Cox-Ingersoll-Ross (CIR) process; and when $\beta = 0$, the asset price follows a standard normal distribution.

The SABR model can explain the phenomenon of the volatility smile due to its flexible parameter structure and its ability to capture changes in volatility under different market conditions. Through the parameter β in the SABR model, the relationship between volatility and the underlying asset price changes can be adjusted, allowing the model to fit different shapes of volatility curves in various market scenarios. For instance, when β approaches 1, the model tends to generate a volatility smile, indicating that volatility increases with larger changes in the forward price, while it decreases when the forward price is more stable. This flexibility enables the SABR model to better capture the volatility smile phenomenon in the

market, thereby enhancing the accuracy and predictive power of option pricing.

3.2 Parameter Estimation

3.2.1 Approximate Solution of SABR

Hagan (2002)[17] adopted singular perturbation techniques to obtain the option prices and derive closed-form solutions for implied volatility.The implied volatility of the option has the following approximate solution:

$$\begin{aligned}
\sigma_B(K,f) &= \frac{\alpha\left(1+\left[\frac{(1-\beta)^2}{24}\frac{\alpha^2}{(fK)^{1-\beta}}+\frac{1}{4}\frac{\rho\beta v\alpha}{(fK)^{\frac{1-\beta}{2}}}+\frac{2-3\rho^2}{24}v^2\right]t_{ex}\right)}{(fK)^{\frac{1-\beta}{2}}\left\{1+\frac{(1-\beta)^2}{24}\log^2(f/K)+\frac{(1-\beta)^4}{1920}\log^4(f/K)\right\}}\cdot\left(\frac{z}{x(z)}\right) \\
z &= \frac{v}{\alpha}(fK)^{\frac{1-\beta}{2}}\log(f/K) \\
x(z) &= \log\left\{\frac{\sqrt{1-2\rho z+z^2}+z-\rho}{1-\rho}\right\}
\end{aligned} \tag{3–3}$$

For at-the-money options, i.e., $K = f$, the volatility calculation formula simplifies as follows:

$$\sigma_{ATM} = \sigma_B(f,f) = \frac{\alpha}{f^{(1-\beta)}}\left\{1+\left[\frac{(1-\beta)^2}{24}\frac{\alpha^2}{f^{2-2\beta}}+\frac{1}{4}\frac{\rho\beta v\alpha}{f^{(1-\beta)}}+\frac{2-3\rho^2}{24}v^2\right]t_{ex}\right\} \tag{3–4}$$

where:f and α respectively represent the initial values of the forward price and volatility, i.e., $F(0) = f$ and $\alpha(0) = \alpha$,t_{ex} represents the expiration time.

3.2.2 SABR Fitting

For each day, I wish to determine the values of SABR parameters based on actual option trading data.The specific procedure involves computing the implied volatility for each strike price using the actual option prices observed during the day. I utilize data for at-the-money and out-of-the-money options. Firstly, the implied volatility for each strike price is calculated using the Black-Scholes formula. Then, the SABR parameters that best fit the implied volatility for that day are determined through the least squares method ($min(\hat{\sigma}(\alpha,\beta,\rho,v)-\sigma_{BS})^2$).

Before this, there is the issue of how to perform least squares fitting. Initially, I used the 'L-BFGS-B' method of the scipy.optimize.minimize function. However, I found that in practical use, if the initial values are not set properly, it may lead to failure to find the optimal

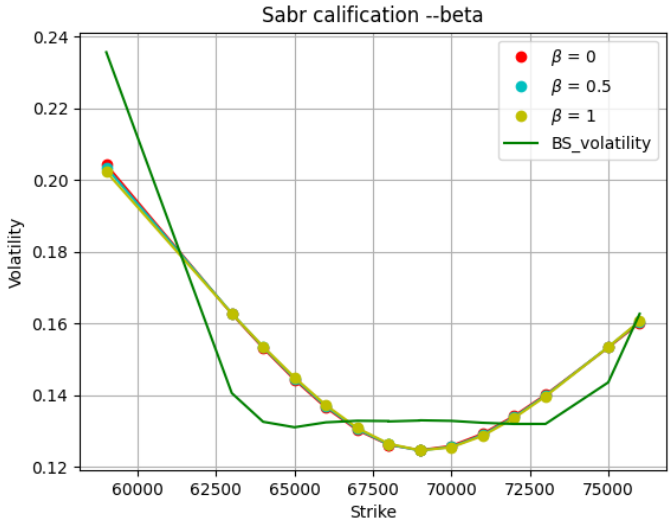


Figure 3–1 Sabr calibration –beta

solution.Hence, I will consider using the simulated annealing algorithm for least squares optimization.

I will examine the impact of each SABR parameter on the implied volatility curve fitted by the SABR model based on actual data.I utilized the trading data of CU2403 dated February 5, 2024, and filtered out trading data with a trading volume accounting for less than 3%. In the estimation of SABR parameters, one issue is the estimation of the parameter beta. A common approach is to fix a value for beta and then estimate the remaining three parameters. Some studies indicate that the impact of beta on the fitting results is minimal. I set beta to 0, 0.5, and 1 respectively to observe the fitted SABR curves as shown in the figure 3–1.

From table 3–1, specific fitting results can be observed, including the indication of the sum of squared residuals. For different values of β, as long as the other three parameters adjust accordingly, the final fitting effect does not vary significantly. Bartlett (B. "Hedging Under SABR Model") noted that the selection of β has limited impact on hedging. However, it is noticeable that when β is set to 0.5, the parameter values are relatively suitable. In practical computations, setting β to 0 or 1 often leads to convergence to local minima. Employing the simulated annealing algorithm with β set to 0.5 yields more stable solutions.

Next, I investigate the impact of the parameter α on the SABR curve. The parameter α represents the initial estimate for volatility. I set the other three parameters as follows:

	SSR	α	ρ	v
$\beta = 0$	0.0025	85.72	-0.0125	0.022
$\beta = 0.5$	0.0026	0.3271	-0.0529	0.0221
$\beta = 1$	0.0027	0.0012	-0.0922	0.0221

Table 3–1 calibration result for different beta

$\beta = 0.5$, $\rho = -0.0529$, and $v = 0.0221$. We vary α as $\alpha = 0.1$, 0.3, and 0.5. The results are shown in the figure 3–2. It is evident that, while keeping the other three parameters constant, variations in the parameter α cause the SABR curve to shift vertically.

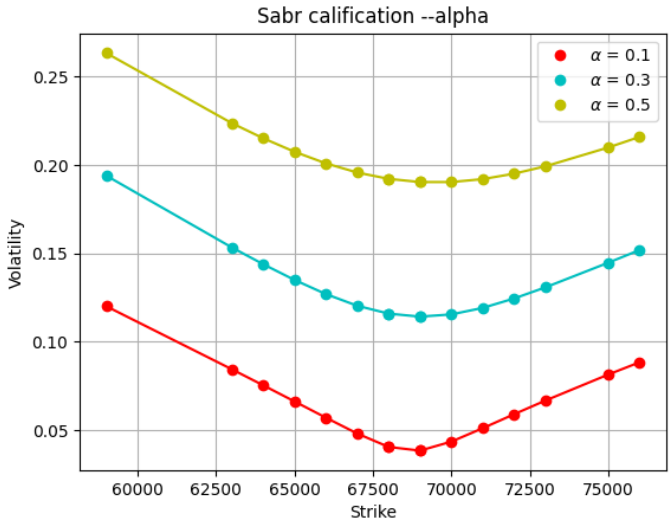


Figure 3–2 Sabr calibration –alpha

Subsequently, I investigated the effect of the parameter ρ on the SABR curve. The parameter ρ represents the correlation between the price of the underlying asset and the stochastic part of volatility. I set the values of the other parameters as $\beta = 0.5$, $v = 0.0221$, and $\alpha = 0.3271$. I varied the parameter ρ to be -0.2, 0, 0.2, and 0.4, respectively. The resulting curves are shown in the figure 3–3. It can be observed from the graph that variations in the parameter ρ cause the SABR curve to shift in a diagonal direction.

Finally, I investigated the effect of the parameter v on the SABR curve. The parameter v represents the volatility of volatility. I set the values of the other three parameters as $\beta = 0.5$, $\rho = -0.0529$, and $\alpha = 0.3271$. I varied the parameter v to be 0.05, 0.15, and 0.3, respectively. The resulting curves are shown in the figure 3–4.It can be observed that the magnitude of

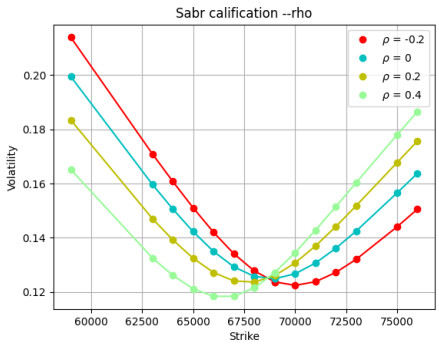


Figure 3–3 Sabr calibration –rho

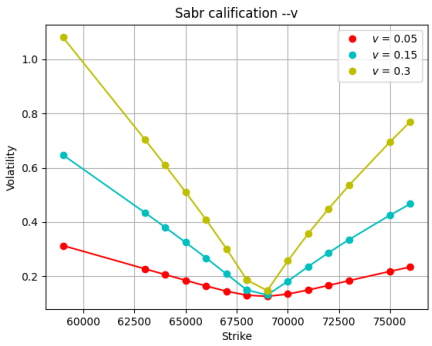


Figure 3–4 Sabr calibration –v

parameter v affects the shape of the SABR curve, with the curvature of the SABR curve decreasing as the parameter v increases.

3.2.3 Weighted SABR Fitting

From the previous figure 2–6, it can be observed that even futures options with certain trading volumes exhibit varying relative trading volumes across different exercise prices. Since the influence of futures options contracts with excessively low trading volumes has already been eliminated, a further refinement is deemed necessary. I argue that futures options contracts with larger trading volumes hold greater importance. Therefore, when fitting the SABR curve, more weight should be assigned to these contracts. Thus, I have modified the optimization objective function from the sum of squared residuals to the sum of squared residuals weighted by trading volume proportions. In other words, the objective function has changed from minimizing $\Sigma(\hat{\sigma}_i(\alpha,\beta,\rho,v)-\sigma_i^{BS})^2$ to minimizing $\Sigma((\hat{\sigma}_i(\alpha,\beta,\rho,v)-\sigma_i^{BS})*w_i)^2$, where w_i represents the trading volume proportion, and $\beta = 0.5$.

On the trading data of CU2403 dated February 5, 2024, I performed a weighted fitting of the SABR model. The resulting fitting plot is shown in figure 3–6. It can be observed that due to the weighting by trading volume, the deviation from the Black-Scholes implied volatility is not significant in the wings, while it is closer to the Black-Scholes implied volatility in the central region.The specific comparison of the fitted parameters and residuals is presented in the table 3–2. From the results, it can be seen that considering the weighting, optimizing the fitting using the new objective function yields a noticeable improvement in residuals. This improvement is reflected mainly in slight shifts and enlargements of the curves, resulting in

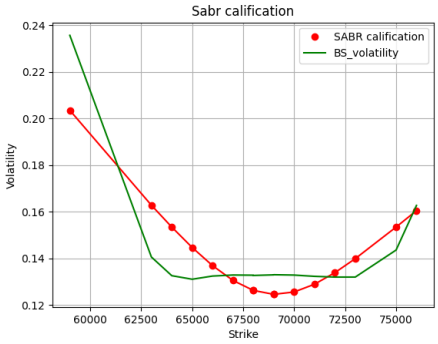


Figure 3–5 Sabr calibration

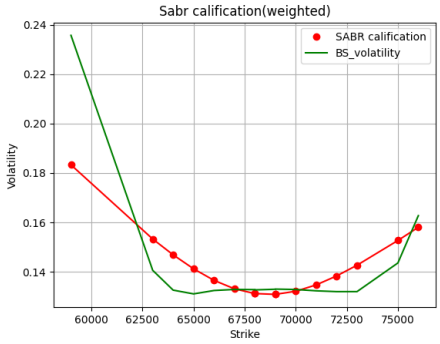


Figure 3–6 Sabr calibration(weighted)

smoother changes in volatility.

	SSR	SSR(weighted)	α	β	ρ	v
not weighted	0.0026	1.59e-05	0.3271	0.5	-0.05290	0.02212
weighted	0.0035	6.10e-06	0.3426	0.5	0.03129	0.01729

Table 3–2 Comparison of weighted and unweighted fitting results

Chapter 4 SABR Model Parameter Prediction for Copper Options

4.1 Data Preparation and Data Analysis

4.1.1 Data Selection and Data Cleaning

I will use copper futures options data from 2021 to April 2024 as an example for prediction.Firstly, the liquidity issue among different futures options on each day will be addressed. For each day within this timeframe, I will select the data of main and second contracts' futures options as the dataset. Secondly, after fixing the futures contract, the liquidity issue among different strike prices on each day will be examined. I will calculate the percentage of trading volume of call (put) options relative to the total trading volume of all call (put) options. Then, I will filter out the trading data where the trading volume proportion is above 3% to obtain valid foundational data.

The specific foundational data extracted includes the trading date, strike price, maturity, option type (flag), transaction price, trading volume, as well as the corresponding forward price of the underlying asset and risk-free interest rate for each day. Due to the difference in maturity between futures options and their underlying futures (futures options settle one month earlier than their underlying futures), I separately computed the time-to-maturity τ_{BS} as the parameter for the Black-Scholes model.The dataset comprises trading data for main and secondary contracts spanning a total of 1046 days, alongside transaction prices for each day' s options contracts corresponding to both main and secondary contracts, where the trading volume for each option exceeds 3% of the total trading volume for all exercise prices.

4.1.2 Black-Scholes Implied Volatility Calculation

I extracted the trading price, exercise price, option type (call, put), underlying futures forward price, and time to expiration (τ_{BS}) from all main and second contract futures options data for 1046 days, resulting in a total of 113764 data points.The specific algorithm involves using the Black-Scholes option pricing equation with the given trading price to iteratively solve for the implied volatility using the Newton-Raphson method.

4.1.3 SABR Calibration

Subsequently, I calibrated the SABR model using option data corresponding to 1046 days of main constract and second constract, totaling 2092 sets of data. The specific approach is to optimize the parameters α, ρ, and v under the condition of given parameter $\beta = 0.5$, such that the weighted sum of squared residuals between the calibrated implied volatilities at various strike prices and the implied volatilities calculated by Black-Scholes, weighted by the trading volumes, is minimized. (Why fitting Black-Scholes implied volatilities instead of actual transaction prices: because Black-Scholes implied volatilities are calculated based on actual prices, in fact, on this day, implied volatilities and transaction prices are equivalent, and then halgon provides an approximate function for calibrating implied volatilities in the paper.)

Figure 4–1 SABR parameters of main constract

My optimization method is simulated annealing algorithm, because in this problem, local optimization algorithms are prone to be trapped in local optima. However, using simulated annealing algorithm introduces new problems. In the case of limited annealing times, it may not converge to the correct solution. Due to the large amount of data, I cannot set too high annealing times. My approach is as follows: based on the previously calculated weighted sum of squared residuals, which is approximately 6×10^{-6}, I will examine the residuals of each optimization result, and if it exceeds 10^{-4}, I will re-optimize, with a maximum of 5 repetitions. If optimization requirements are not met in 5 attempts, the optimization result is

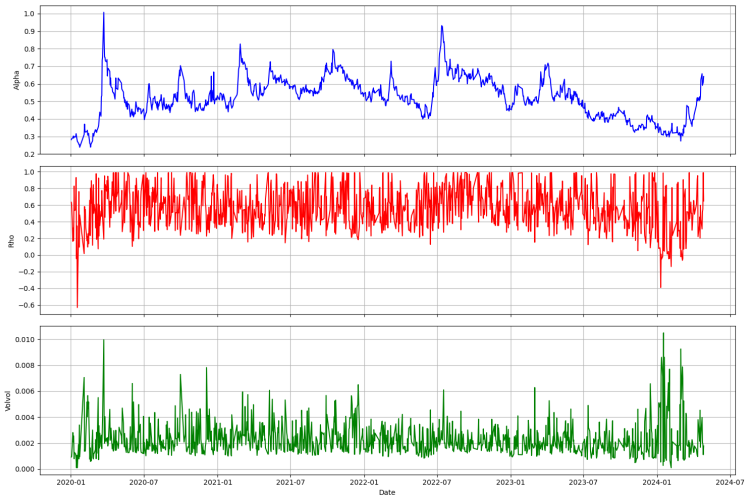


Figure 4–2 SABR parameters of second constract

recorded as ”nan”. This effectively avoids the problem of poor results in SABR calibration optimization.The time series plot of SABR parameters for the main contract is shown in Figure 4–1, while the time series plot of SABR parameters for the second contract is shown in Figure 4–2.

From the time series Figure 4–1, it can be observed that there are some extreme values in the optimized parameters, and these extreme values are concentrated in the same time period. Secondly, regarding the time series trends of the parameters, it is evident that the trend of parameter α is relatively pronounced, exhibiting a cyclic repetition process, while the time trends of the other parameters are not as directly observable.

4.1.4 Preliminary data analysis

I conducted a preliminary analysis on the basic data of the main contract and second contract during this time period. Firstly, I am concerned about the changes in the forward prices of copper futures because, if the goal is to predict the prices of futures options, it is necessary to determine the prices based on the Black-Scholes pricing model. In addition to establishing a predictive model for implied volatility, it is also necessary to forecast the changes in the underlying futures' forward prices to predict the prices of futures options accurately.

I plotted the trend of the forward prices of the main contract and second contract's underlying futures over time in the Figure 4–3. By observing the general trend of price changes, I

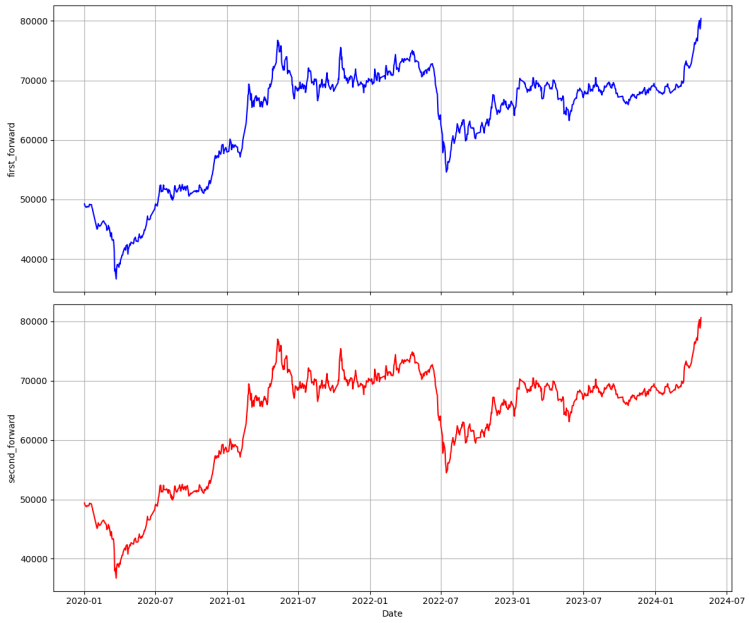


Figure 4–3 Forward price of copper futures

aim to extract useful information from past price movements to forecast future price changes. Referring back to the extreme value points in Figure 4–1, I noticed that they correspond to the periods of significant changes in forward prices (As for why there were no extreme situations in Figure 4–2, it may be attributed to probability; repeating the calculations several times might not yield similar extreme situations as in Figure 4–1, but since they occurred, they require an explanation). This indicates that drastic changes in forward prices contribute to corresponding changes in SABR parameters (especially α).

When considering the changes in forward prices of the underlying futures, another crucial parameter is the time remaining until the expiration date. This is because the main contract and second contract are not fixed futures; they generally last for a period and then transition to the next futures contract (for example, in February 2023, it might be CU2304, and in March, it transitions to CU2305). Therefore, the forward prices of the main contract and second contract do not represent the forward prices of a fixed-term option. I believe that future changes in forward prices may follow the patterns of past price changes, implying that the time remaining until the expiration date, denoted as τ, may contain significant information.

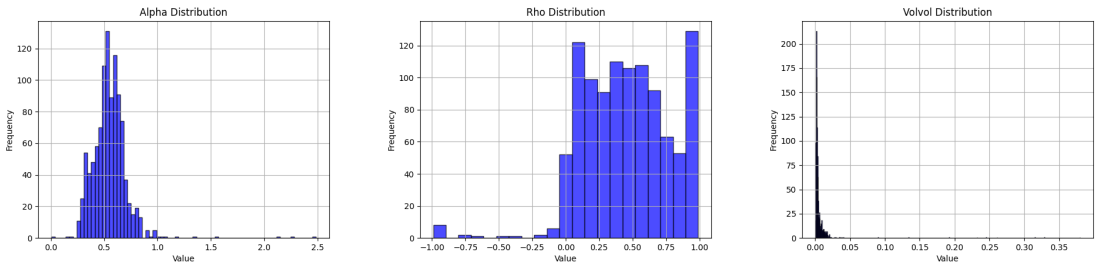
Next, I conducted data analysis on the SABR parameters, namely α, ρ, and v. Firstly, I calculated their means and standard deviations, as shown in Table 4–1 (I use α_1 to rep-

resent the parameters of the main contract and α_2 to represent the parameters of the second contract.). It can be observed that, compared to the main contract, there is a relatively large difference in parameter v for the second contract. From the standard deviations of the three parameters, it is evident that the parameter ρ exhibits the most significant variation, followed by α, while the variation in parameter v is minimal. Subsequently, I plotted histograms to visualize the distribution of the data.

	α_1	ρ_1	v_1	α_2	ρ_2	v_2
mean	0.549873	0.455522	0.006337	0.521618	0.565853	0.002307
standard deviation	0.170598	0.330980	0.023919	0.115508	0.251722	0.001288

Table 4–1 Mean and standard deviation of SABR parameters

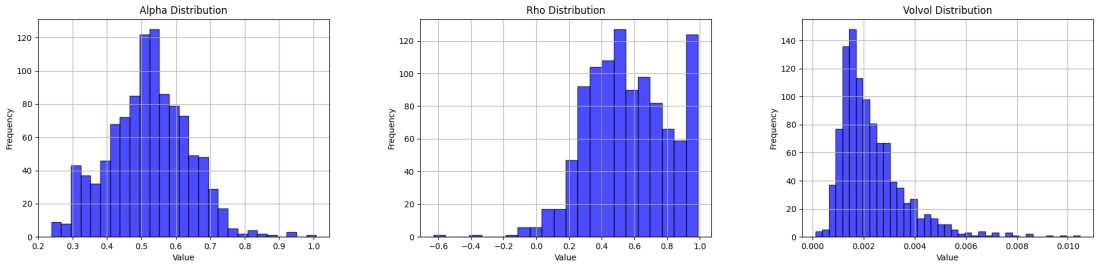
From the histogram 4–4 of the data distribution for the main contract, it can be observed that there are a few outliers in the α parameter. I will remove the main contract data where α is less than 0.2 or greater than 1.2 to ensure the data's validity. Similarly, I will remove the main contract data where v is greater than 0.025. The distribution figure 4–5 of parameters for the second contract appears relatively normal, with no excessively strange values appearing. In the outliers of the main contract data, once one parameter takes an extreme value, the values of other parameters tend to be extreme as well. Therefore, as requested, only 16 data points need to be removed in the end.



Alpha Distribution **Rho Distribution** **Volvol Distribution**

Figure 4–4 Parameter distribution of main contract

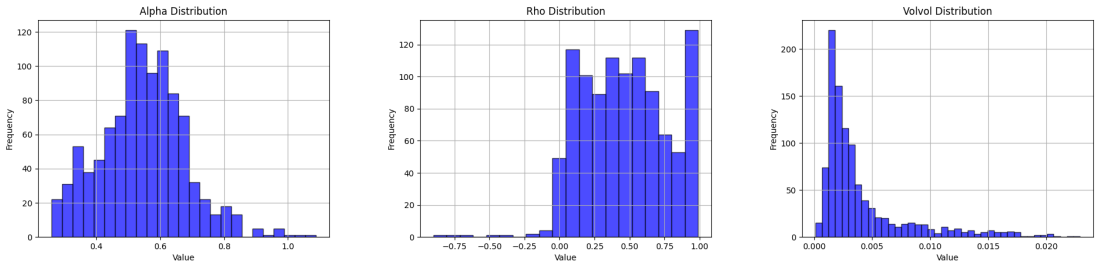
I obtained the SABR parameters for the main contract after cleaning the data, and then plotted the distribution histogram 4–6 and time series graph 4–7. From the distribution histogram, it can be seen that there is a strong similarity between the main contract and the



Alpha Distribution **Rho Distribution** **Volvol Distribution**

Figure 4–5 Parameter distribution of second contract

second contract, and the time series trends of the SABR parameters are also similar. I will use the cleaned data to construct models for prediction.



Alpha Distribution **Rho Distribution** **Volvol Distribution**

Figure 4–6 Parameter distribution of main contract(cleaned)

4.2 Predictive Model

4.2.1 ARIMA Model

1. **MA (Moving Average) Model:**

The moving average model of order q, denoted as $\mathrm{MA}(q)$, is defined as

$$Z_t = \theta_0 + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}, \tag{4–1}$$

where q is a natural number, and $\{a_t\} \sim \mathrm{WN}\left(0, \sigma_a^2\right)$ **ReferencesForMA**.

2. **AR (Auto Regression) Model:**

An autoregressive model of order p, denoted as $\mathrm{AR}(p)$, satisfies

$$Z_t = \theta_0 + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \cdots + \phi_p Z_{t-p} + a_t, \tag{4–2}$$

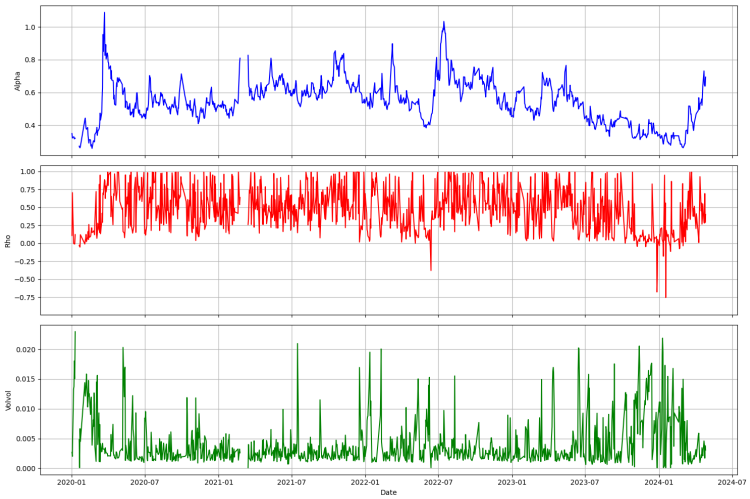


Figure 4–7 SABR parameters of main constract(cleaned)

where p is a natural number, ϕ_i are real numbers, and $\{a_t\} \sim \mathrm{WN}\left(0, \sigma_a^2\right)$**ReferencesForAR**.

3. **ARMA Model:**

If Z_t satisfies

$$\begin{aligned} Z_t &= \theta_0 + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \cdots + \phi_p Z_{t-p} \\ &\quad + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}, \end{aligned} \tag{4–3}$$

then Z_t is called an autoregressive moving average of orders p and q, denoted as ARMA(p, q)[19].

A time series $\{Z_t\}$ is considered stationary if

- $\mathbb{E}(Z_t)$ is constant;
- $\mathrm{var}(Z_t)$ is constant;
- $\mathrm{cov}(Z_t, Z_{t-k})$ is independent of t for all $k = 1, 2, \ldots$.

The models of time series analysis are built upon the foundation of stationary time series. Non-stationary data often result in "spurious regression," where two variables with no causal relationship exhibit high correlation. For instance, in a time series, two time points that originally have no autocorrelation may show correlation. If a time series is not stationary, it needs to be transformed into a stationary one through differencing.

4. **ARIMA Model:**

Consider the following two points:

1. By differencing, non-stationary time series may become stationary. 2. The ARMA model is the most widely studied stationary time series model.

Combining the above two points, we obtain the integrated autoregressive-moving average (ARIMA) model[19].

4.2.2 GARCH Model

In traditional econometric models, the variance of the disturbance term is assumed to be constant. However, many economic time series exhibit volatility clustering, where the assumption of constant variance is inappropriate.

1. **ARCH Model:**
   - The disturbances $\{a_t\}$ in the asset return series are serially uncorrelated but not independent;
   - The lack of independence in $\{a_t\}$ can be described by a simple quadratic function of its lagged values.

   Specifically, an ARCH (m) model is given by:

$$\begin{aligned} a_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_m a_{t-m}^2 \\ \alpha_0 &> 0; \forall i > 0, \alpha_i \geqslant 0 \end{aligned} \tag{4–4}$$

   where ε_t is an independent identically distributed (iid) sequence of random variables with mean 0 and variance 1. It is typically assumed to follow a standard normal distribution. σ_t^2 represents conditional heteroskedasticity.

2. **GARCH Model:**

   Bollerslev (1986) proposed a generalized form known as the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model[12].

   Let $a_t = r_t - u_t$ denote the innovation at time t. If a_t satisfies the following equation:

$$\begin{aligned} a_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2 \\ \alpha_0 &> 0; \forall i > 0, \alpha_i \geqslant 0, \beta_j \geqslant 0, (\alpha_i + \beta_j) < 1 \end{aligned} \tag{4–5}$$

   where ε_t is an independent identically distributed (iid) sequence of random variables with mean 0 and variance 1. σ_t^2 represents conditional heteroskedasticity. Then a_t is said to follow a GARCH (m, s) model.

4.2.3 LSTM

Recurrent Neural Networks (RNNs) are a type of neural network used for processing sequential data, capable of handling data with sequential changes[20]. For instance, the meaning of a word may vary depending on the preceding context, a scenario where RNNs excel in resolving such issues. The specific process of an RNN is illustrated in Figure 4–8; when passing parameters, not only the current time step x_t is inputted but also the previous time step h_{t-1}. h_t is calculated as

$$h_t = \sigma\left(x_t \times w_{xt} + h_{t-1} \times w_{ht} + b\right) \tag{4–6}$$

This means that compared to traditional Deep Neural Networks (DNNs), RNNs can utilize past inputted data to construct the neural network, making it suitable for processing time series data.

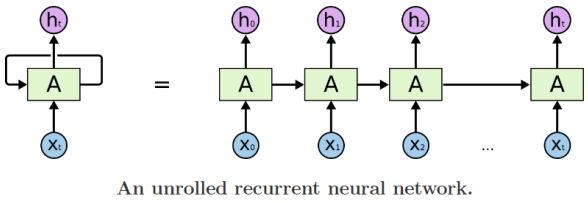


Figure 4–8 RNN

However, RNNs encounter issues such as the inability to handle long-term dependencies[21]. Long Short-Term Memory (LSTM) networks, introduced by Hochreiter and Schmidhuber[22], do not suffer from such issues and are particularly effective for tasks that require memory of distant events.

LSTMs possess a more complex structure than RNNs and incorporate three gates: the forget gate (f), input gate (i), and output gate (o). The specific operations are as follows:

$$\begin{aligned}
f_t &= \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \\
i_t &= \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \\
\tilde{C}_t &= \tanh\left(W_C \cdot [h_{t-1}, x_t] + b_C\right) \\
C_t &= f_t * C_{t-1} + i_t * \tilde{C}_t \\
o_t &= \sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right) \\
h_t &= o_t * \tanh\left(C_t\right)
\end{aligned} \tag{4–7}$$

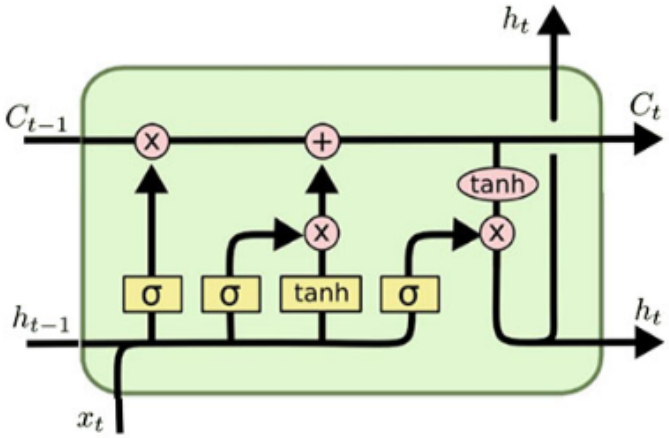


Figure 4–9 LSTM

Contrasting with RNNs, LSTM introduces a modification to h_t; it selectively forgets certain information while utilizing C_t to convey information, thus reducing the risk of gradient vanishing.

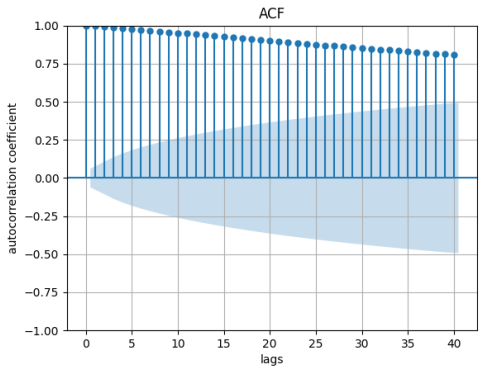
4.3 Forecast Methods

Here, from a practical perspective, I will mainly choose the model from two directions. Firstly, the first direction is time series modeling. I will construct time series models to predict changes in forward prices and SABR parameters. Specifically, this involves either ARIMA models or GARCH models, depending on the actual situation. The second direction is to build neural network models, specifically LSTM. I will select the last 20 data points to compare the predictive performance of the two. The reason is that training LSTM models can be challenging. In practical applications, I may only update the model once a month, corresponding to approximately 20 trading days. Therefore, I choose to compare the predictive performance of the two models over 20 days.

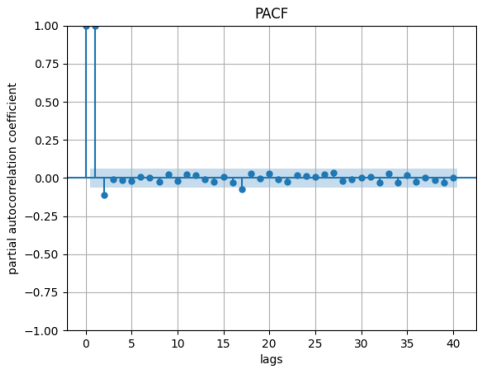
4.3.1 Time Series Models ARIMA

As I need to predict four variables, I use the model of forward prices as an example. Firstly, I observe the time series plot of forward prices, where I notice an upward trend in forward prices. Then, I compute the ACF and PACF plots for the forward price data. From

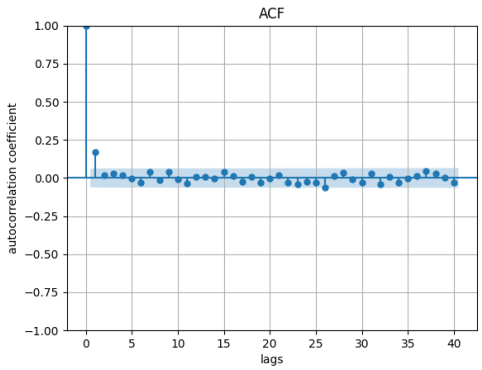
the ACF Figure 4–10, it is evident that the data is non-stationary. Therefore, I perform differencing on the data and plot the ACF and PACF plots again. The differenced ACF and PACF plots exhibit clear truncation features. Hence, I consider an ARIMA model with d=1. Based on the ACF and PACF Figure 4–10, I determine the range of p and q to be approximately 0-3. Within this range, I fit all possible ARIMA models and select the one with the smallest AIC as the final fitted model, which is ARIMA(0,1,1).



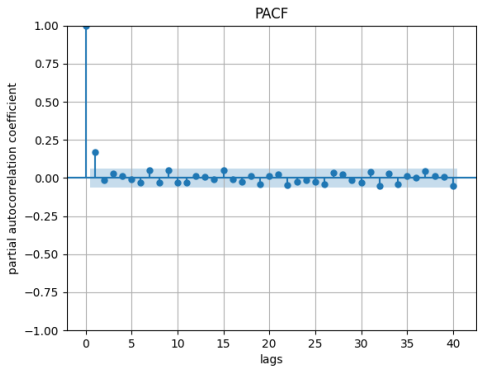
ARIMA-ACF



ARIMA-PACF



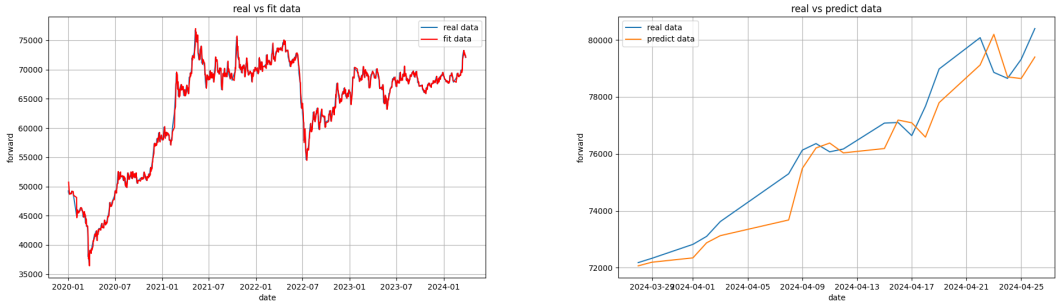
ARIMA-ACF-d1



ARIMA-PACF-d1

Figure 4–10 ARIMA : ACF and PACF

Regarding the prediction part, if I directly forecast the next 20 data points, the predictions generated in the previous steps will be used for iterative forecasting, leading to error accumulation and convergence of predictions to a single value. ARIMA performs poorly for long data series. However, by obtaining the parameters from the ARIMA model and incrementally forecasting future data using actual data, i.e., predicting only the next day's data and updating the model daily, such prediction Figure 4–11 can be obtained.



real vs fit data **real vs predict data**

Figure 4–11 ARIMA Predict

I use the same approach to construct an ARIMA time series model for α_t, ρ_t, and v_t, and predict the data for the last 20 days, storing my forecast results.As shown in the figure 4–12, I have also plotted the fitting image of SABR parameters using ARIMA.

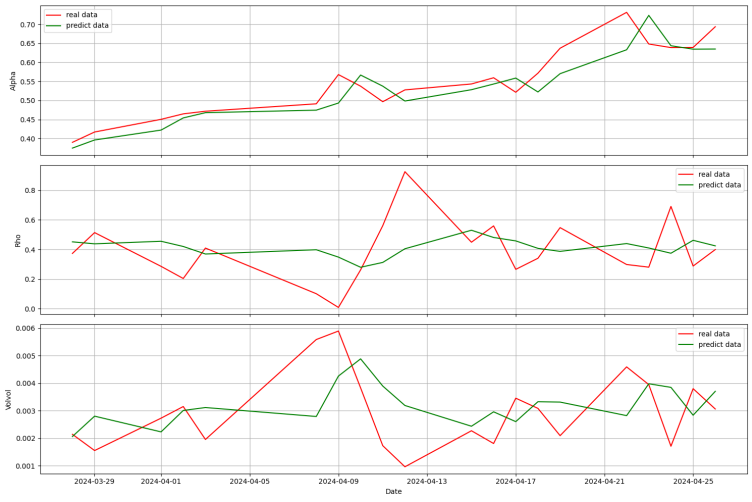


Figure 4–12 ARIMA-real vs predict SABR parameters

Upon careful observation of the fitted images, it is not difficult to notice that both the predicted images for the parameter α and the forward price F exhibit a lagging attribute, closely resembling the previous value. This also reflects a drawback of time series models, namely their monotonic nature, which makes it difficult to discern the distribution patterns of complex data. Particularly for univariate time series models, it is challenging to capture the relationships between data. In this practical scenario, it is evident that there are relationships between SABR parameters. Thus, a univariate time series model may not be entirely suitable.

4.3.2 Time Series Models GARCH

When fitting an ARIMA model to the forward, the residuals are assumed to be Gaussian white noise series. Thus, when we fit the data with an ARIMA model, we need to perform an LB test on the estimated residual sequence to determine if it is Gaussian white noise. If not, it suggests that the ARIMA model may not be suitable for the sample. Here, we use the forward as an example. For the residuals of the model fitting, I conduct an Ljung-Box test. The Ljung-Box test is a statistical method used to examine whether there is autocorrelation in the residual sequence of a time series. It is based on the statistic calculated from the auto-correlation function (ACF) of the time series and is commonly used to assess the fitting effect of time series models and check whether the residuals of the model meet the assumptions. The constructed statistic is:

$$Q = n(n+2)\sum_{k=1}^{h}\frac{\hat{\rho}_k^2}{n-k}$$

where n is the sample size, $\hat{\rho}_k^2$ is the sample autocorrelation coefficient of lag k, and the statistic follows a chi-square distribution with h degrees of freedom. Given the significance level α , the rejection region is $Q > \chi^2_{1-\alpha,h}$. Accepting the null hypothesis implies that the original sequence is a white noise sequence, otherwise, it suggests the existence of correlation in the sequence. The results obtained are as follows:

lb_stat	lb_pvalue
0.587097	0.443544
0.587281	0.745544
0.606694	0.894898
0.606996	0.962283
0.613732	0.987366
0.653859	0.995433
0.681134	0.998477
0.682405	0.999569
0.740028	0.999839
0.754166	0.999954

According to the results of the Ljung-Box test, after considering a series of lag orders, the residual sequence of my model does not exhibit significant autocorrelation. This indicates that my model has captured the information in the time series relatively well.

4.3.3 LSTM

When fitting the LSTM model, there are several factors to consider. Firstly, the time step, which determines how much data to input for prediction. I have decided to base this on 20 data points, corresponding to one month's time. I hope that by using data from one month, the model can learn relevant patterns and make accurate predictions for the future. Secondly, the input and output parts need to consider which data are useful. I have already calculated sequences for the SABR parameters and forward prices. I believe the remaining time until the expiration date is also crucial information. Additionally, the trading volume of options and corresponding futures on the same day might contain important data. Next is the output data. Naturally, to make accurate predictions of option prices, the output data consists of the forward price and SABR parameters.

During the training process of the model, I have allocated the last 20 data points to the test set, with the remaining data serving as the training set. The input parameters are: the forward price of the underlying futures, the SABR parameters α, ρ, ν, the trading volume of the underlying futures, and the trading volume representing the options, as well as the time remaining until the exercise date of the option. The output parameters are: the forward price of the underlying futures, the SABR parameters α, ρ, ν. Then, the input is a 20-length data sequence, and the output is a single-length data point, which means that I will use one month's trading data to predict the trading data for that day. After adjusting the neural network parameters, I found that the neural network model performs quite well under these parameters, which are as follows and include three layers:

LSTM Layer (Long Short-Term Memory Network): The output shape of this layer is (None, 20, 24), indicating that the output is a two-dimensional array where the first dimension is None (representing variable length), the second dimension is 20 (representing the time steps), and the third dimension is 24 (representing the number of features at each time step). The number of parameters is 3,168.

Dropout Layer: The output shape of this layer is the same as that of the LSTM layer,

(None, 20, 24), with no additional parameters. It is used to randomly drop out some neurons during the training process to prevent overfitting.

TimeDistributed Layer: The output shape of this layer is (None, 20, 4), indicating that the output is a two-dimensional array where the first dimension is None (representing variable length), the second dimension is 20 (representing the time steps), and the third dimension is 4 (representing the number of output features at each time step). The number of parameters is 100.

The total number of parameters is 3,268, all of which are trainable, with no non-trainable parameters.

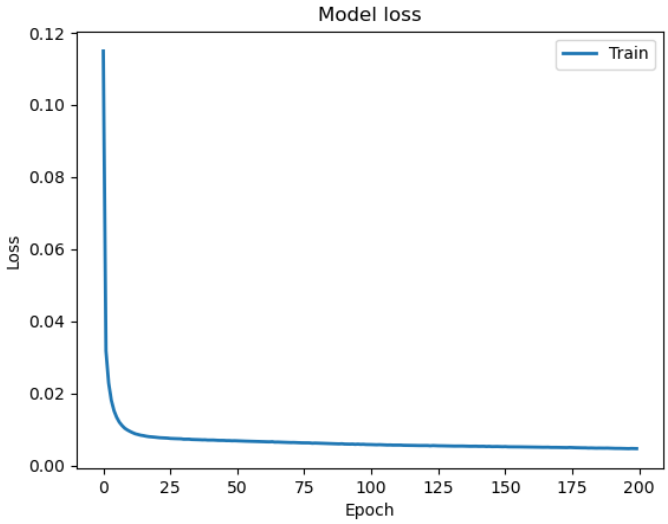


Figure 4–13 LSTM learning

The training history plot is shown in the following Figure 4–13 ,

The final fitting results are shown in the Figure 4–14. After 200 rounds of training, the model's accuracy has reached 0.85. It can be observed that there is a slight deviation in the prediction of the SABR parameters, however, slight fluctuations in the SABR parameters are not particularly sensitive to the final results. There is a significant improvement in the prediction of the forward price of the underlying futures.

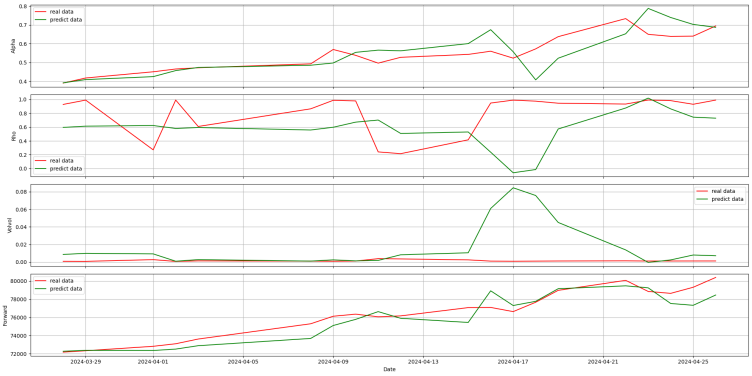


Figure 4–14 LSTM:real vs predict

4.4 Forecast Result

4.4.1 Parameter Prediction Results

In terms of practical application, I plan to forecast the data for one year following January of the year 2023. From a practical standpoint, as the data is continuously updated, the model also needs to be updated regularly. Then, considering the time series model ARIMA is easily updatable, I can update the parameters of the ARIMA model on a daily basis. However, training an LSTM model is more complex, so it can only be updated once a month.

According to the method I described, I have used both ARIMA and LSTM approaches to predict the variations in SABR parameters and forward prices over the course of a year. Of course, I have performed calculations for the main contract and the second contract, as well as for the main and second contracts without weighting by trading volume. Furthermore, I have compared the LSTM method with the ARIMA method. Below is the analysis based on the data of the main contract with SABR parameters corrected by trading volume weighting.

Firstly, observing the forward price, I have plotted the actual data and the forecasted data after January 1, 2023, on the graph 4–15. Additionally, I have calculated their respective Mean Squared Errors (MSE) as shown in the table 4–2. Initially, it can be seen from the graph that compared to ARIMA, the fluctuations in the LSTM predicted values are more pronounced, making it more susceptible to producing more extreme values. In contrast, the ARIMA forecast results are mostly minor variations between the previous few data points. From the MSE, it is evident that in terms of predicting forward prices, the ARIMA model performs better.

data	ARIMA	LSTM
Forward	215165.27188219552	522602.7750101382

Table 4–2 MSE for Forward Price Predictions

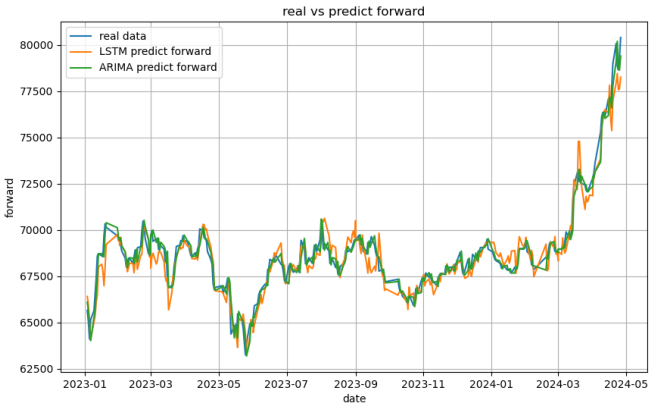


Figure 4–15 LSTM vs ARIMA forward

Following the analysis, I have plotted a comparison graph 4–16 for the predictions of the SABR parameter α and provided a table 4–3 for the Mean Squared Errors (MSE). For the prediction of the SABR parameter α, the results from the LSTM model do not match up to those of the ARIMA model. The primary reason for this is that the variation in the SABR parameter is more sensitive to randomness, indicating that a significant portion of the fluctuations may be stochastic in nature, with substantial impact. Consequently, ARIMA, which is less susceptible to such disturbances, relies only on the preceding values for its predictions. In contrast, the LSTM model is more heavily influenced by randomness, as evidenced by some of the more extreme forecasts it has produced. This is also related to the limited amount of data used for training the models.

data	ARIMA	LSTM
alpha	0.001541188736604399	0.002221059771364767

Table 4–3 MSE for alpha Price Predictions

For the prediction of the SABR parameter ρ, I have similarly plotted comparative graphs 4–17 and presented the respective Mean Squared Errors (MSE) in a table 4–4. Consistently, it is observed that the predictive performance of the LSTM model does not surpass that of the

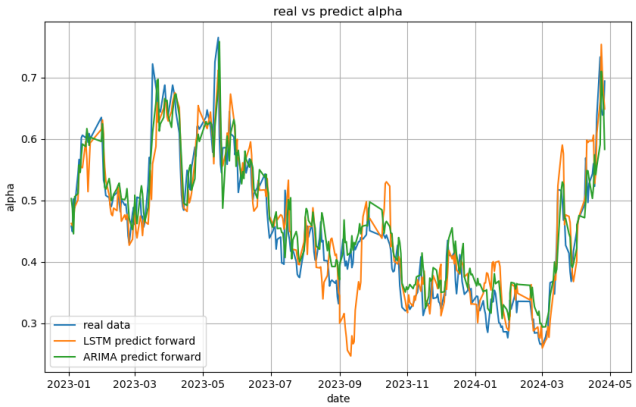


Figure 4–16 LSTM vs ARIMA alpha

ARIMA model. A significant factor contributing to this lies in the inherent characteristics of the SABR parameter ρ, which is the correlation coefficient between two Brownian motions within the SABR model, bounded between -1 and 1. The LSTM model has produced predictions exceeding this threshold, indicating an incomplete grasp of the parameter's constraints by the model. It is suggested that modifications to the LSTM model, possibly incorporating constraints reflective of the parameter's nature, could potentially enhance its predictive accuracy.

data	ARIMA	LSTM
rho	0.11335579775162628	0.14753469504582117

Table 4–4 MSE for rho Price Predictions

For the prediction of the SABR parameter v, which represents the volatility of volatility in the SABR model and is strictly positive, I have also plotted comparative graphs 4–18 and presented the respective Mean Squared Errors (MSE) in a table 4–5. Similarly, it is observed that the LSTM model's predictive performance does not match that of the ARIMA model, and in this aspect, the LSTM's performance is particularly poor. This issue stems from the same challenge faced in predicting ρ; despite its stronger learning capabilities, the LSTM model has not effectively captured certain regularities from the limited data. During the calibration process, the parameter v tends to be calibrated very close to zero, which, like the correlation coefficient ρ approaching 1, can lead to fitting outcomes that exceed the parameter's valid

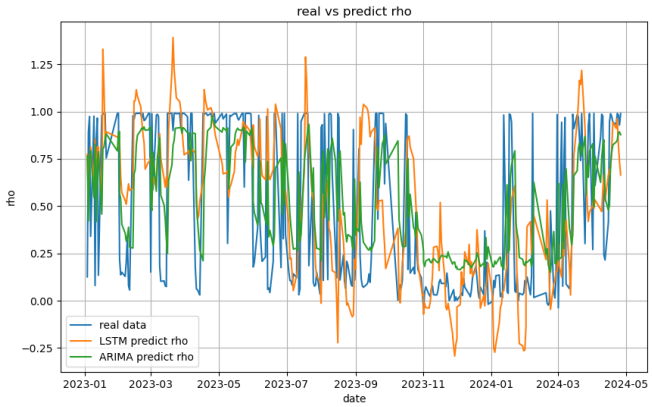


Figure 4–17 LSTM vs ARIMA rho

range when no constraints are imposed. The LSTM model, without such constraints, is prone to producing fits that fall outside the permissible bounds.

data	ARIMA	LSTM
rho	1.6371554021117275e-05	0.00013930700534505776

Table 4–5 MSE for volvol Price Predictions

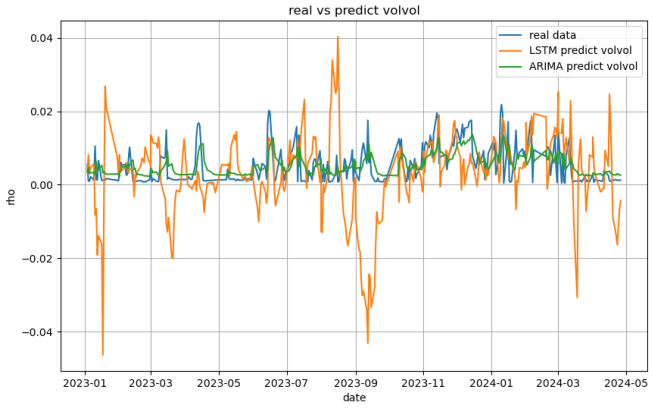


Figure 4–18 LSTM vs ARIMA volvol

Considering the aforementioned issues, I have made slight modifications to the LSTM network architecture. Specifically, I have implemented a custom output layer by adding a new Dense layer to the model structure and defining a custom activation function using a

Lambda layer. In this layer, I have constrained the first output (α) and the third output (v) to be greater than 0.0001. This is because, in SABR calibration, the parameter α represents the volatility and must be a positive number. Additionally, I have limited the second output (ρ) to a range between -0.99 and 0.99. This is due to the fact that, in SABR calibration, the parameter ρ represents the correlation coefficient, which should have a value within the range of -1 to 1.

After making the aforementioned adjustments, I obtained new forecast results. I plotted the forecast graphs and calculated the Mean Squared Error (MSE), which is recorded in the table 4–6. It can be observed that, although the modified LSTM model has shown some improvement in the prediction of parameters α and v, it still does not outperform the ARIMA model. The main reasons for this are twofold: 1. The issue of data volume: I only utilized two years of daily frequency data to train the LSTM model, which is far from meeting the basic requirements for the LSTM to perform optimally. 2. The randomness of data fluctuations: Since the SABR parameters are derived from SABR calibration, the degree of stochastic volatility is relatively high, making it challenging to model and predict. 3. The parameter tuning of the LSTM: Although I have endeavored to select the best possible parameters, there may still be room for improvement.

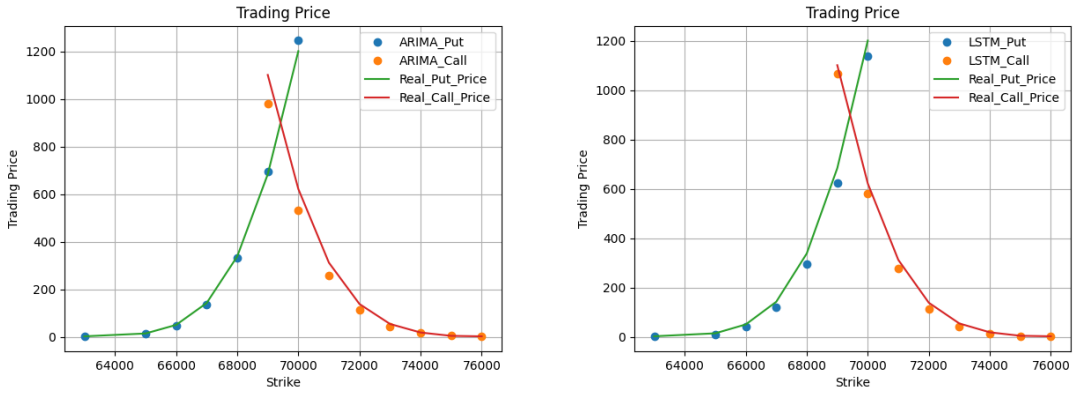
Parameter	ARIMA MSE	LSTM MSE	LSTM New MSE
Forward Price	215165.27	522602.7750	639499.464
α	0.001541	0.0022210	0.001900550
ρ	0.113355	0.1475346	0.18988954
v (Volvol)	1.637155e-05	0.000139307	5.4428969e-05

Table 4–6 MSE for Forward Price, α, ρ, and v Predictions

4.4.2 Price Prediction Results

My ultimate goal is to forecast the price changes of futures options. The prediction of parameters is merely one step in this process. Subsequently, I intend to use the predicted parameter results to forecast the variations in option prices. This implies that I will employ the SABR calibration to predict the implied volatility under various strike prices, and subsequently utilize this implied volatility along with the Black-Scholes formula to predict the option prices.

Here, we observe two sets of comparison charts. I have chosen the predicted data of the main contract (CU2403) on February 1, 2024. The first set of charts 4–19 compares the prediction performance of ARIMA and LSTM. The horizontal axis of the chart represents the exercise price, and the vertical axis represents the trading price of options. I have plotted both the actual trading prices and the predicted trading prices. From the chart, it can be seen that ARIMA outperforms LSTM in price prediction, which is consistent with the conclusion drawn in the previous section. That is, in the case of insufficient data, LSTM cannot predict the data well. Of course, this also illustrates the second fact, that is, from the overall fitting effect, according to the parameters calibrated by SABR, through parameter prediction, calculating the implied volatility at each exercise price, and finally, based on the logic of predicting the forward price for the pricing of futures options, there is no problem. If more accurate predictions of SABR parameters and forward prices can be achieved, more accurate prediction results will be obtained.

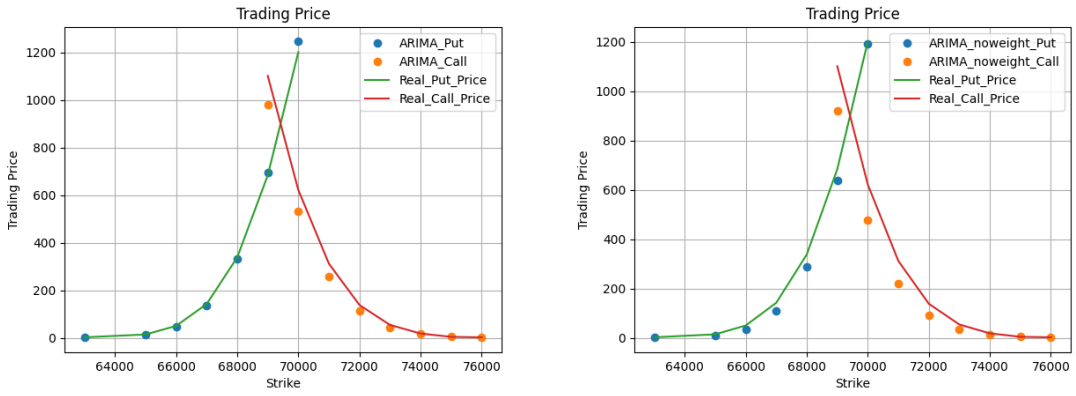


Price Forcast Arima **Price Forcast LSTM**

Figure 4–19 ARIMA vs LSTM Price Predict

Next, I compared the price prediction charts of the SABR parameters calibrated with and without volume-weighted calibration. From the charts 5–4, it can be seen that, similar to the previous images, the horizontal axis represents the exercise price, and the vertical axis represents the trading price. Additionally, call options and put options are distinguished by different colors. Firstly, it should be noted that when the exercise price is close to the forward price, the trading volume is higher, corresponding to the middle position in the charts. For deep in-the-money and deep out-of-the-money options, the trading volume is low. There-

fore, for call options, the trading volume concentrates on the middle-right position, while for put options, the trading volume concentrates on the middle-left position. Then, from the comparison of the two charts, it is evident that the predicted values of the SABR parameters calibrated with volume weighting perform better at these points with higher trading volumes. Indeed, I have achieved the goal of more accurately predicting the prices of futures options by considering trading liquidity.



Price Forcast Arima **Price Forcast Arima unweighted**

Figure 4–20 Weighted vs Unweighted Price Predict

Chapter 5 Trading Strategy and Backtest Results

5.1 Trading Strategy

Based on my predictive results, I will develop a simple yet effective trading strategy that reflects the main point of the article (the importance of trading liquidity). I will determine whether to open long or short positions based on the trend of predicted option prices. The specific steps are as follows:

- Calculate the rate of change of predicted prices (i.e., predicted returns) and set upper and lower thresholds. When the predicted return exceeds the upper threshold, if there are no positions, open a long position; otherwise, open a short position. If already holding long positions, no trading is done. If holding short positions, close them first before opening new positions.
- For futures options at different strike prices, determine the position size based on the trading volume ratio. Specifically, if the ratio is below 3%, no trading is done; for ratios between 3% and 5%, the position size is set to 0.1; for ratios between 5% and 10%, the position size is set to 0.3; for ratios between 10% and 15%, the position size is set to 0.5; and for ratios above 15%, the position size is set to 1.
- Set the maximum holding period to 10 days.
- If the futures on the corresponding day are not the main contract or the second contract, stop predicting and close any existing positions directly (because each option is only part of the time the main contract or the second contract, and the data is not collected for prediction during the remaining time).
- Set stop-loss and take-profit conditions: once the price falls by more than 5% compared to the purchase price, trigger the stop-loss condition and force liquidation; once the price falls by more than 5% compared to the highest price after purchase, trigger the dynamic take-profit condition and force liquidation.

I simply apply this strategy to conduct trades. Because my purpose is not to find the optimal strategy, I just need to horizontally compare the impact of trading liquidity studied and the differences in the prediction methods I set. So I try to make the strategy as simple as possible for experimentation. My specific approach is to calculate the returns relative

to today's real prices based on the trading prices I predicted. Once the returns are greater than 0, I take a long position; if they are less than 0, I take a short position. As for position management, I continue the specifics of my strategy, setting positions based on the proportion of trading volume for different strike prices of the same option. If the proportion is less than 3%, no position is taken. Positions are set to 0.1 for proportions between 3% and 5%, 0.3 for proportions between 5% and 10%, 0.5 for proportions between 10% and 15%, and 1 for proportions above 15%. I believe this approach is in line with investment logic and also fits the main points of this paper. As for the overall returns for the day, I will calculate the weighted average of the options held based on the computed positions. Then, I only hold the positions for the day I make the prediction. For simplicity, I did not set transaction fees and did not optimize strategy parameters.

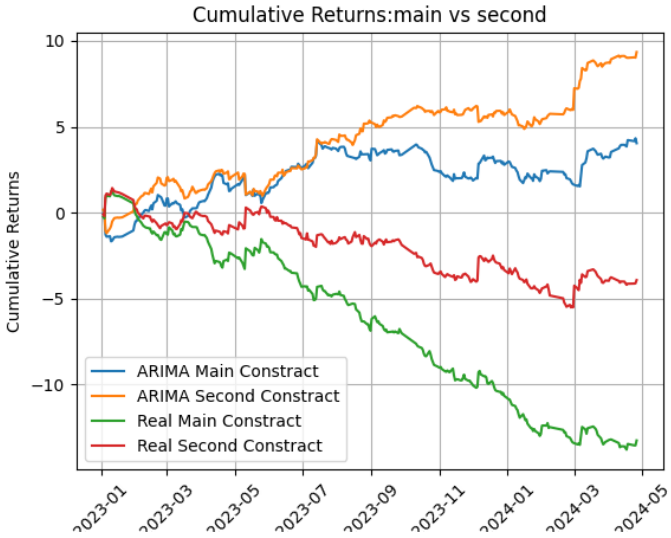


Figure 5–1 Cumulative Returns: main vs second

First, let's look at the overall effect. This figure 5–1 shows the cumulative returns obtained by executing the strategy based on the predictions obtained from ARIMA, including data for both the main contract and the secondary contract. Because in a day, the options and futures trading corresponding to the main contract and the secondary contract generally account for over 70% of all options and futures trading for the same underlying asset, as explained in Chapter 2; it can be seen that both achieve decent cumulative returns, with higher cumulative returns for the secondary contract, indicating the prediction performance for the

secondary contract. Then, the graph also shows the variation curve of real returns, which has also undergone position adjustments. It can be observed that the cumulative returns obtained by combining prediction with strategy are significantly higher than those of no operation, indicating the effectiveness of my predictions. Furthermore, it can also be observed that perhaps the original real returns of the main contract were not high, resulting in the cumulative returns of the main contract being lower than those of the secondary contract. In fact, the excess returns of the two datasets are almost the same. This table 5–1 lists some parameters calculated for the daily returns for both contracts. From the data in the table, it can be seen that the main contract exhibits stability with lower annualized returns and volatility, while the secondary contract shows more activity with higher annualized returns and volatility, as well as a higher Sharpe ratio, demonstrating superior performance.

Parameter	Main Contract	Second Contract
Annualized Return (%)	3.18	7.35
Annualized Volatility (%)	3.47	3.40
Sharpe Ratio	0.91	2.16
Maximum Drawdown (%)	-2.66	-1.51
Return to Drawdown Ratio	1.20	4.88
Win Rate	50.9%	55.7%
Profit/Loss Ratio	0.98	1.05

Table 5–1 Cumulative Returns：main vs second

In this figure 5–2, I compared the performance differences between ARIMA and LSTM predictions on the main contract. It can be observed that, compared to ARIMA, LSTM exhibits greater volatility. Additionally, in certain time periods, LSTM outperforms ARIMA. The differences in positions between the two also appear significant, indicating that LSTM has learned something that ARIMA cannot recognize, and has also achieved good performance.

Next, I will compare the data obtained from predicting using SABR parameters calculated with and without volume weighting. I chose the results of the LSTM method on the secondary contract for a more intuitive comparison.In the Figure 5–3, it can be seen that the cumulative returns without weighting are far inferior to those obtained with weighting. There are two main reasons for this. Firstly, when calibrating SABR parameters without weight-

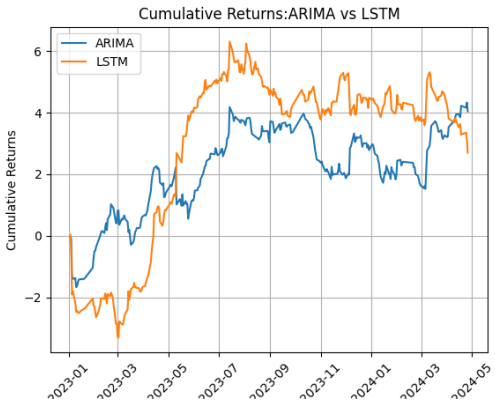


Figure 5–2 Cumulative Returns：ARIMA vs LSTM

Figure 5–3 Cumulative Returns：Weighted vs Unweighted

Figure 5–4 Weighted vs Unweighted Price Predict

ing, the fitting of SABR parameters in areas with high trading volume is reduced, which is contradictory to the strategy, as the trading strategy assigns more position weight to options with higher trading volume. Therefore, under this strategy, the daily weighted returns are inevitably affected. The second reason, as mentioned in Chapter 2, is that for options with poor liquidity, price changes are more difficult to capture. Due to the phenomenon of volatility smile, their volatility is higher, making them more difficult to trade at normal prices. To eliminate the influence of the first reason, I also calculated the returns using unweighted positions for the predictions without weighting. It can be seen that the cumulative returns are indeed higher than those with weighted positions, but still far inferior to those predicted with weighted calibrated SABR parameters.

Chapter 6 Summary

6.1 Key Findings

Firstly, this paper provides an overview of option pricing methods, ranging from the Black-Scholes option pricing model to stochastic volatility models, with specific elucidation on the Heston model and the SABR model. The focus is on elucidating the SABR model. Real trading data is employed to conduct a detailed investigation of option pricing models. Implied volatility indicators are computed from option trading prices and strike prices, uncovering the phenomenon of volatility smiles. To further explicate the volatility smile phenomenon, stochastic volatility models are introduced, followed by SABR model calibration using real data. This process involves an examination of the fundamental principles of SABR calibration, the specific roles of each parameter in the calibration process, and the resultant impact on the final calibrated curve. In this segment, the primary viewpoint and key conclusion are presented: only trading information with volume holds value. Additionally, specific methodologies and rationale are proposed. The approach entails: 1) selecting main and secondary contracts as fundamental data for the study, and 2) trading volume-weighted SABR calibration, where enhancements to the original SABR calibration process are made. Subsequently, an illustration is provided, depicting the disparities between SABR calibration curves with and without volume weighting.

Next, the primary objective of option pricing is strategy formulation. Backtesting of strategies is performed to observe specific returns. In line with the main viewpoint, it is anticipated that the experimental phenomenon will indicate higher returns from SABR calibration with volume weighting compared to those without. In the experimental section, copper futures options serve as an example. Firstly, main and secondary contract data for copper futures options post-2020 are employed to calculate implied volatility indicators using the Black-Scholes equation, followed by SABR calibration to obtain daily SABR parameter indicators. Secondly, models are constructed to model historical data and predict future data. ARIMA and LSTM models are utilized to model historical data, with continuous refinement throughout the comparative analysis of both models. Thirdly, once models are established, future data is predicted. To align closely with reality and derive conclusive findings, predic-

tions are made for data post-2023. Given the abundance of options data in real life (though only copper is utilized here), model updates are performed every month to capture the latest volatility trends. Lastly, based on the predicted results, a simple trading strategy considering trading volume proportion is devised and executed in accordance with the prediction results. Ultimately, the conclusions are validated, affirming that SABR calibration considering trading volume proportion leads to higher and more stable returns.

6.2 Research Outlook

Before discussing the research outlook, let me address the existing limitations of this study. While theoretically, conducting volume-weighted SABR calibration poses no theoretical challenges, several deficiencies remain. I will outline these shortcomings in order of significance:

1. **Flaws in Prediction Models**: In the experimental section, I utilized both the ARIMA and LSTM models. However, the ARIMA model's simplicity restricts its ability to capture nuanced information, while the LSTM model's performance is hindered by insufficient training data. Consequently, the predictive efficacy of these models is compromised, thereby impacting subsequent backtesting procedures.
2. **Volatility in SABR Parameters**: During the discussion on SABR calibration and the examination of parameter influences, it was observed that the β parameter exerts minimal impact on the final calibrated curve. Hence, the β parameter was fixed at 0.5. Nonetheless, challenges persist. SABR parameters exhibit high sensitivity to variations in the SABR curve, and observed transaction prices may fluctuate around theoretical values. Consequently, SABR parameters derived from actual data may deviate significantly from theoretical parameters, rendering accurate predictions with the established model more challenging.
3. **Data Volume Issues during Validation**: While comprehensive validation encompassing all options for strategy backtesting and analysis would yield more compelling conclusions, computational and temporal constraints necessitated an illustrative example using copper options.

Moving forward, there exists a promising avenue to address the aforementioned issues, notably through the implementation of Kalman filtering to simultaneously predict and cali-

brate data. This approach effectively mitigates real value fluctuations, thereby substantially enhancing model performance. However, given current limitations in theoretical knowledge, this is deemed as future work. A preliminary exposition on Kalman filtering, based on personal understanding and theoretical diagrams, is presented in the initial subsection of this chapter.

Furthermore, in future prospects, enhanced experimental validation via Kalman filtering could spur theoretical advancements. For instance, while SABR parameter calibration employed linear weighting in this study, exploring alternative weighting methodologies and incorporating additional data for weighting represent potential avenues for further exploration.

Additionally, refinement of trading strategies remains an open frontier. While a rudimentary trading strategy considering trading liquidity was proposed herein, it remains distant from a mature, comprehensive strategy. Although validating the conclusions of this paper suffices, significant room for improvement exists for achieving greater excess returns in option quantification trading.

6.2.1 Kalman Filter

The Kalman filter, proposed by Rudolf E. Kalman in the 1960s[23], is an efficient recursive filter that can estimate the state of a dynamic system from a series of incomplete and noisy measurements. The Kalman filter considers the joint distribution of measurements at different time points to generate estimates for unknown variables, making it more accurate than estimation methods based solely on individual measurements. The basic working principle of the Kalman filter is as follows:

The data available for analysis includes time series of forward prices and SABR parameters, while the known data comprises trading dates, days to expiration, strike prices, and risk-free interest rates. The objective is to predict the trading prices of specific futures options at various strike prices.

Due to inherent volatility in transaction prices, accurate transaction prices are inevitably unpredictable; I believe the core principle of the Kalman filter is to recognize that data inherently contains noise, rendering any predictions and actual observations inaccurate. I aim to eliminate noise from this noisy data, extract the trend component of prices, and obtain filtered data. Based on the filtered data, I will then predict new data.

The Kalman filter primarily consists of three processes—prediction, correction, and updating.

The first process is the prediction process, where the next system state at time step t+1 is predicted based on the current estimated state (not the true value), expressed as:

$$x_{t+1}^{\text{predicted}} = Ax_t^{\text{estimated}} + Bu$$

$$P_{\text{predicted}} = A \cdot P_{\text{estimated}} \cdot A^T + Q$$

where $x_{t+1}^{\text{predicted}}$ is the predicted state, $x_t^{\text{estimated}}$ is the current estimate of the state, A is the state transition matrix, B is the control input matrix, u is the control input, $P_{\text{predicted}}$ is the predicted estimation error covariance, $P_{\text{estimated}}$ is the current estimation error covariance, and Q is the process noise covariance. Here, Q accounts for the noise present regardless of the theoretical model used, which introduces discrepancies between true and predicted values.

The second process is the correction process, where the actual value at time t+1 can be observed, but it is the estimated value with added noise. To obtain the estimated value without noise, considering the measurement process, I might not directly measure x_{t+1}^{real}, but indirectly measure z_{t+1}^{real}. A typical example is predicting the distance between two objects, where the time for light to propagate between them is measured instead. Then, I multiply it by a transformation matrix. The formula is:

$$z_{t+1}^{\text{real}} = Hx_{t+1}^{\text{real}} + \varepsilon$$

where z_{t+1}^{real} is the measured real value at time t+1, x_{t+1}^{real} is the actual value we want to obtain, H is the transformation matrix created based on the relationship between the two values. Considering measurement errors, in financial data, this can be understood as true data being estimated data plus noise, with bias from the noise of true data. Assume the covariance matrix of ε is R.

The third process is the updating process. How do we obtain $x_{t+1}^{\text{estimated}}$ from $x_{t+1}^{\text{predicted}}$ and x_{t+1}^{real}? Since both $x_{t+1}^{\text{predicted}}$ and x_{t+1}^{real} include noise and are inaccurate, I will decide based on the covariance matrices of both. I will be closer to the one with smaller covariance matrix because I believe its estimation is more accurate and noise is smaller. The formulas are:

$$K = P_{\text{predicted}} \cdot H^T \cdot (H \cdot P_{\text{predicted}} \cdot H^T + R)^{-1}$$

$$x_{t+1}^{\text{estimated}} = x_{t+1}^{\text{predicted}} + K \cdot (z_{t+1}^{\text{real}} - H \cdot x_{t+1}^{\text{predicted}})$$

$$P_{\text{estimated}} = (I - K \cdot H) \cdot P_{\text{predicted}}$$

where K is the Kalman gain, incorporating covariance of observed and predicted values. $H \cdot P_{\text{predicted}} \cdot H^T + R$ is the covariance matrix of the observed value z_{t+1}^{real}, and it's noticeable that when one side's covariance matrix is larger, the estimation tends to lean towards the other side.

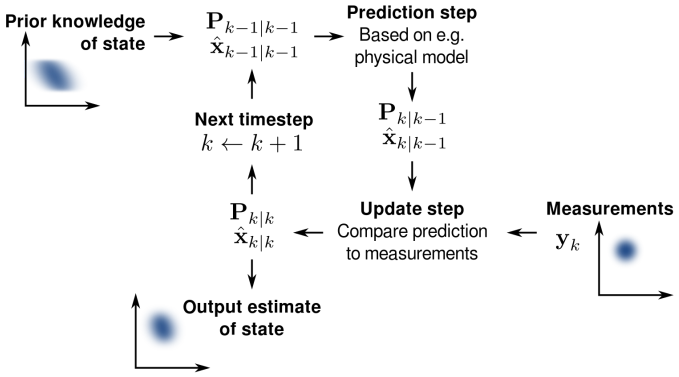


Figure 6–1 Basic concept of Kalman filtering

From the Figure 6–1, one can observe the fundamental principles of the Kalman filter.

Bibliography

[1] BLACK F, SCHOLES M. The Pricing of Options and Corporate Liabilities[J]. Journal of Political Economy, 1973, 81(3): 637-654. DOI: 10.1086/260062.

[2] MERTON R C. Theory of Rational Option Pricing[J]. Bell Journal of Economics and Management Science, 1973, 4(1): 141-183. DOI: 10.2307/3003143.

[3] BLACK F. The Pricing of Commodity Contracts[J]. Journal of Financial Economics, 1976, 3(1-2): 167-179. DOI: 10.1016/0304-405X(76)90024-6.

[4] RUBINSTEIN M. 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[J]. The Journal of Finance, 1985, 40(2): 455-480.

[5] HULL J C. Options, Futures, and Other Derivatives[J]. 2006.

[6] YAN S. Jump risk, stock returns, and slope of implied volatility smile[J]. Journal of Financial Economics, 2011, 99(1): 216-233.

[7] GATHERAL J. The Volatility Surface: A Practitioner's Guide[J]. 2006.

[8] DUMAS B, FLEMING J, WHALEY R E. Implied Volatility Functions: Empirical Tests[J]. The Journal of Finance, 1998, 53(6): 2059-2106.

[9] CLARK P K. A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices [J]. Econometrica, 1973, 41(1): 135-155.

[10] TAUCHEN G, PITTS M. The Price Variability-Volume Relationship on Speculative Markets[J]. Econometrica, 1983, 51(2): 485-505.

[11] ENGLE R F. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation[J]. Econometrica, 1982, 50(4): 987-1007.

[12] BOLLERSLEV T. Generalized Autoregressive Conditional Heteroskedasticity[J]. Journal of Econometrics, 1986, 31(3): 307-327.

[13] HULL J C, WHITE A. The Pricing of Options on Assets with Stochastic Volatilities[J]. The Journal of Finance, 1987, 42(2): 281-300.

[14] TAYLOR S. Modeling Financial Time Series with STABLE(1) Noise[J]. The Journal of the American Statistical Association, 1987, 82(398): 1125-1137.

[15] SCOTT L O. Option Pricing when the Variance Changes Randomly[J]. The American Journal of Agricultural Economics, 1987, 69(2): 547-557.

[16] WIGGINS J B. Option Valuation under Stochastic Volatility: Evidence that the Market Overreacts [J]. Journal of Financial Economics, 1987, 19(2): 351-372.

[17] HAGAN P S, KUMAR D, LESNIEWSKI A S, et al. Managing Smile Risk[J]. Wilmott Magazine, 2002, 2(8): 84-108.

[18] HESTON S L. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options[J]. The Review of Financial Studies, 1993, 6(2): 327-343.

[19] BOX G E P, JENKINS G M. Time Series Analysis: Forecasting and Control[M]. San Francisco: Holden-Day, 1976.

[20] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors[J]. Nature, 1986, 323(6088): 533-536.

[21] BENGIO Y, SIMARD P, FRASCONI P. Learning long-term dependencies with gradient descent is difficult[J]. IEEE transactions on neural networks, 1994, 5(2): 157-166.

[22] HOCHREITER S, SCHMIDHUBER J. Long Short-Term Memory[J]. Neural computation, 1997, 9(8): 1735-1780.

[23] KALMAN R E. A New Approach to Linear Filtering and Prediction Problems[J]. Journal of Basic Engineering, 1960, 82(Series D): 35-45.

Acknowledgements

As four years of university swiftly pass by, I find myself at the culmination of my undergraduate journey. Reflecting on the past four years, vivid scenes flash before my eyes, transporting me back to the classroom where Professor Han Dong taught probability theory, where Professor Wang Cheng lectured on mathematical statistics... The surroundings may blur with time, but the knowledge etched on the blackboard remains crystal clear.

Here, I realize there are countless individuals I am deeply grateful for. I thank all my professors for imparting knowledge, my class advisor for guiding me through life's paths, and my counselor for their supportive presence in my daily life. Special gratitude goes to Professor Samuel Drapeau, who wholeheartedly supervised my thesis. From selecting the topic, gathering data, coding, to completing the paper, he offered immense support at every step, guiding me through the exploration of this thesis.

Furthermore, I extend heartfelt thanks to my parents, relatives, and loved ones for their unwavering support. Their everyday conversations have been my greatest source of emotional sustenance during moments of confusion and despair. It is their financial backing, care in daily life, and encouragement that have enabled me to successfully complete my studies and move forward to the next phase of my life.

Lastly, I express sincere gratitude to the busy reviewers who took the time to evaluate my work and to all the teachers who provided valuable feedback during the initial and mid-term defense of my thesis.