LUND UNIVERSITY
School of Economics and Management

Master's Programme in Data Analytics and Business Economics

Option Pricing with Machine Learning: An Application of Shanghai Crude Oil

DABN 01 Master's Thesis

Shuiqing Pan

Supervisor: Joakim Westerlund

May 2024

Abstract

Options, as complex derivative assets, play a crucial role in real life for hedging risks and speculating profits. Their high volatility and the challenging nature of price predictions make them not only essential tools for financial markets but also a compelling research topic for scholars. Building on previous studies, this thesis focuses on Shanghai Crude Oil options, offering insights into their pricing dynamics. In addition to traditional pricing features, such as strike price, time to maturity, and volatility, this thesis explores the impact of macroeconomic factors on crude oil prices to determine whether these factors can also influence option pricing. By employing machine learning algorithms, the performances of these models are compared with the Binomial Tree model. The findings show that XGBoost model outperforms the benchmark in both cases of Call and Put options, while other models' performance shows less robustness.

Key words: Options Pricing, Shanghai Crude Oil, Machine Learning, Macroeconomic Factors

Contents

List of Figures

List of Tables

List of Abbreviations

ANN	Artificial Neural Network
CNY	Chinese Yuan
EPU	Economic Policy Uncertainty Index
fx_close	Closing Price of Exchange rate
GPR	Geopolitical Risk Index
GRU	Gated Recurrent Units
LSTM	Long Short-Term Memory
K	Strike Price
RF	Random Forest
RNN	Recurrent Neural Network
SC	Shanghai Crude Oil
Settle_fut	Settle Price of SC Future
Settle_opt	Settle Price of SC Option
SVM	Support Vector Machine
T	Time to Maturity
USD	United States Dollars
WTI	West Texas Intermediate
wti_close	Closing Price of WTI Crude Oil Future
XGBoost	eXtreme Gradient Boosting

1 Introduction

Derivatives are important financial instruments for market participants to hedge and speculate. The dynamic and high volatility nature of derivatives also presents a compelling research area, particularly in developing effective pricing strategies. Among various derivative products, options stand out due to their complexity and high-dimensional features, making accurate and fast pricing models essential for market participants.

The Black-Scholes (BS) model introduced by Black and Scholes (1973) provides a foundational understanding of options pricing. This model is designed for European-style options and has some constraints due to the ideal assumptions. Later, the revised version and improvements have been conducted by scholars over the years (Merton, 1973; Cox et al., 1979; Engle, 1982; Heston, 1993; Hull and White, 1987) and the improved model has widely adopted in the industry (Li, 2022). American options, which can be exercised at any time prior to and including the expiration date, present additional complexities over the European types. This flexibility makes the valuation of American options more challenging and more sophisticated, where the Binomial Tree model by Cox et al. (1979), the Trinomial tree model by Boyle (1986) and Finite Difference Methods and Least Squares Monte Carlo simulation (Longstaff and Schwartz, 2001) can be used to address the early exercise problems.

With the development of artificial intelligence and significant advancements in computational technology, the potential to enhance these models through machine learning has opened new avenues for research, specifically in improving predictive accuracy and optimizing algorithmic trading strategies. The integration of large datasets and sophisticated programming techniques can substantially contribute to improving the precision and speed of pricing models, offering substantial benefits to traders and speculators. The earliest study is from Hutchinson et al. (1994), who utilized the Artificial Neural Networks (ANN) models to pricing options, and proved this neural network worked effectively in the case of S&P 500 futures options. Subsequent research has focused on deep learning techniques for more accurate price prediction, with Long Short-Term Memory (LSTM) networks showing particularly strong performance (Bennell and Sutcliffe, 2004; Culkin and Das, 2017; Buehler et al., 2019; Zhang and Huang, 2021; Liu and Zhang, 2023).

While prior studies have investigated more from the cases of the index options, this thesis aims to narrow down to a specific commodity and analyze Shanghai Crude Oil (SC) options. Choosing SC option as a research target can be particularly interesting and beneficial for several reasons. First, SC options are American-type, adding complexity and making them suitable for testing machine learning techniques. Additionally, China plays a crucial role as one of the largest consumers and importers of crude oil globally (Yi et al., 2021). Then from the perspective of derivatives asset, since its launch, SC options have seen significant growth in trading volume and the future price is highly correlated with the global market, including West Texas Intermediate (WTI) Crude Oil and Brent Crude Oil (Shanghai International Energy Exchange (INE), 2024). Therefore, studying these options provides insights into Chinese market dynamics regarding oil consumption and trading.

For the factors influencing oil prices, although the Organization of Petroleum Exporting Countries (OPEC) occupies a dominant role. By establishing production quotas for its member countries, OPEC can influence oil prices through supply control, although its control over the market is not absolute, other factors including the rise of alternative energy sources, the varying economic needs of its member countries, and oil production from non-OPEC countries also contribute to the pricing volatility (Colgan, 2014). Other than OPEC's influences, research on oil prices is a comprehensive and detailed field of study. Hamilton (2003), Bennell and Sutcliffe (2004) and Kilian (2009) demonstrated that economic growth has a positive influence on oil price. Golub (1983), Amano and Van Norden (1998) and Jo (2012) confirmed the fluctuations in the US dollar (USD) exchange rate and central banks' monetary policies, which influence interest rates and liquidity, also affect oil prices. Uncertainty factors including the economic and geopolitical may increase supply disruptions, then affect global oil prices (Yi et al., 2021).

Building on the prior research basis, deep learning techniques such as ANN and LSTM have demonstrated robust predictive capabilities. This thesis aims to utilize machine learning techniques to explore the option pricing model, incorporating factors from traditional pricing theories and macroeconomic conditions to broaden the understanding of how global economic events and policy changes influence the Chinese oil markets. The benchmark model is the Binomial Tree model without macroeconomic factors. The algorithms to be tested include Random Forest, Support Vector Machine, eXtreme Gradient Boosting, Artificial Neural Network, Long Short-Term Memory, and Gate Recurrent Unit. It seeks to identify and refine the most effective computational models for enhancing predictive accuracy in this SC option case.

The outline of the thesis is as follows:

- **Chapter 2: Literature Review** – This chapter discusses the previous research of both traditional and machine learning-based option pricing methods, highlighting key developments and findings in the field. The traditional models are mainly demonstrated from the models designed for American options.
- **Chapter 3: Data** – This chapter focuses on data utilized in this thesis, detailing the resources from which the data was sourced, and describing the preprocessing procedures.
- **Chapter 4: Methodology** – This chapter details the benchmark model and machine learning techniques used in this thesis, including model selection, and evaluation criteria.
- **Chapter 5: Results and Discussion** – This chapter presents the findings from the empirical analyses, comparing the performance of machine learning models against traditional approaches.

- **Chapter 6: Conclusion** – This chapter summarizes the research outcomes, and discusses the implications for practitioners.

2 Literature Review

The evolution of option pricing models has been significantly influenced by advancements in both financial theory and computational technologies. From the nature of reliance on predefined assumptions, these models have been classified into two principal categories: parametric methods and non-parametric models. In this part, previous papers are analyzed with such structures, in addition to the specific crude oil price analysis.

2.1 Traditional models

Traditional option pricing models are built on predefined mathematical equations and assumptions about market behaviors and the statistical properties of asset prices. Black and Scholes (1973) introduced the BS with seven ideal assumptions, which revolutionized financial markets with its analytical approach. Originally developed to price European options, which are exercisable only at maturity, the BS model presumes constant volatility—a condition often unrealistic in actual market settings. This critical limitation was addressed in the same year by Merton (1973) through introducing stochastic volatility by a jump-diffusion term. This extension of the BS model, now known as the BSM model, accommodates more realistic market conditions and suits options with features and exercise rights that are more diverse.

Brennan and Schwartz (1977) demonstrated how to use the Finite difference method (FDM) to price American options. FDM is a numerical approach that utilizes partial differential equations to predict the option's value. The method applies boundary conditions through discretizing the equations on a grid in both price and time. Therefore, it is appropriate for American options, considering the payoff condition and early exercise feature.

Boyle (1977) pioneered to adopt Monte Carlo simulation for option pricing. This method employs a random sampling approach to estimate the expected payoff of an option, which is then discounted to the present using the risk-free rate. Monte Carlo simulation excels in handling complex derivative structures that are not easily addressed by traditional models, such as path-dependent and American options which require the flexibility to model multiple sources and sequences of randomness. The method's inherent flexibility allows it to manage exotic options and calculate various valuation metrics, including the Greeks, which quantify the sensitivities of an option's price to key underlying factors. While Monte Carlo simulations provide significant adaptability and precision, they do come with a higher computational burden. They require extensive simulations to capture the full range of potential future paths for the underlying asset's price, significantly increasing computational time and resource usage, especially as the complexity of the financial instrument increases.

Later, Cox et al. (1979) introduced a Binomial Tree model for American option pricing by allowing the early exercise of options before expiration. As the name indicates, each node in the model represents a possible price of the financial assets at a given time, and each node branches into two possible future nodes, reflecting the potential upward or downward movements of the price. This branching structure enables the model to simulate the various paths that the stock price might take over the life of the option, which is suitable for American options as it may be exercised at any time before the expiration date. Although the model is simple and easy to apply, the computational costs become expensive as the number of steps rise. Developed from the binomial model, Boyle (1986) extends the tree model by introducing a third possible state: stay the same at each step. The trinomial model allows for better adjustments to the probabilities of different price movements, making it better suited to capture the skewness and kurtosis of price distributions.

2.2 Machine learning models

Emerging with the advances in computer science, particularly in the areas of machine learning and data processing capabilities, non-parametric models do not assume a previously specified functional form and instead rely on data to model complexities in market dynamics. Techniques such as neural networks, decision trees, and kernel regression have been employed to estimate the underlying distributions without the constraints of parametric assumptions. These models are particularly valuable in environments where the underlying asset characteristics are not adequately described by traditional parametric models. Researchers have employed non-parametric methods to investigate pricing techniques for various types of options, primarily focusing on European and American options.

The study by Hutchinson et al. (1994) was among the first to apply non-parametric techniques to pricing options. The primary goal of Hutchinson et al. (1994) was to determine whether neural networks could effectively learn the patterns and complexities in financial markets without relying on the strict assumptions required by traditional economic models. By analyzing options on S&P 500 futures, it has been concluded that ANN models demonstrate superior accuracy and computational efficiency compared to the BS model. This landmark study challenged the traditional models which relying on strict assumptions, indicating that the data-driven approaches could perform robust and accurate predictions.

Following Hutchinson's work, a number of researchers have further studied this topic. Yao et al. (2000) examined the volatility structures of the market, enhancing the predictive power of these models under various market conditions. Meanwhile, Gençay and Salih (2003) focused on refining regularization techniques to improve model stability and performance. In a comparative study, Bennell and Sutcliffe (2004) demonstrated that neural network models could effectively compete with and sometimes outperform various configurations of the BS model in predicting market movements. More recently, Culkin and Das (2017) replicated and extended Hutchinson's initial findings by employing a feed-forward neural network to model

BS pricing mechanisms, confirming the robustness and applicability of neural networks in financial modeling.

Given the sequential nature of financial time-series data, RNN—and specifically LSTM —have been adopted to capture temporal dependencies that other models might overlook. Buehler et al. (2019) explored LSTM for developing hedging strategies, demonstrating their effectiveness in managing risk in dynamic trading environments. Zhang and Huang (2021) applied LSTM networks across various market datasets, contrasting their performance with traditional models under different conditions. Their research concluded that LSTM generally outperforms traditional models in scenarios characterized by low to medium market volatility, specific levels of moneyness, and within certain risk thresholds. These findings underscore the potential of advanced RNN to transform financial prediction and risk management strategies, offering a significant upgrade over more rigid, less adaptive models. A newly released study by Liu and Zhang (2023) examined a LSTM model with realized skewness. For the dataset, they used the ETF50 option in China and concluded that the model with realized skewness performs better than the benchmark models, including classical and other machine learning methods in all metrics.

2.3 Oil price influencer

Extensive researches have been conducted on oil prices, since it is a critical and highly valued commodity. Previous studies have explored various macroeconomic factors influencing oil prices from multiple perspectives.

Supply and Demand: OPEC, a cartel consisting of 13 of the world's major oil-exporting nations, aims to manage the supply of oil to set the price on the world market. The organization has its dominant role, but external factors such as the rise of alternative energy sources, the varying economic needs of its member countries, and oil production from non-OPEC countries (Colgan, 2014). Fattouh and Sen (2016) examined OPEC's influence considering non-OPEC oil production and shifting global demand dynamics, concluding that OPEC's role as a price maker is not absolute.

Economic Growth: Hamilton (2003) and Bennell and Sutcliffe (2004) found that oil prices tend to rise in periods of robust economic growth, reflecting higher energy demand. Kilian (2009) further distinguished this relationship between supply and demand shocks, emphasizing how global economic growth pushes up oil demand and prices.

Exchange Rates: Golub (1983) established a strong correlation between exchange rate fluctuations and oil prices in the application of USD movement to global oil prices. This relationship was further supported by Amano and Van Norden (1998), who discussed the implication of on monetary policy on oil-importing economies.

Interest Rate: Basher et al. (2012) highlighted that higher interest rates tend to depress the oil prices by influencing the speculators' behaviors. Jo (2012) suggested that volatile interest rates have a continuous effect on oil price, which in turn affect the inflation and broader economy.

Economic Policy Uncertainty (EPU) and Geopolitical Risk (GPR): Aloui et al. (2016) and Antonakakis et al. (2017) found that higher EPU or increased market volatility, leading to higher level of risk aversion among investors and affected the crude oil price. This sentiment was extended by Balcilar et al. (2017), suggesting how increased GPR negatively affects oil prices. Wei et al. (2017) and Cunado et al. (2020) also analyzed these factors from behavioral finance perspective, confirming high levels of EPU and GPR make investors more cautious about the crude oil futures, leading to worse liquidity. For the long-term effect, Hu et al. (2020) found that investors required more risk premium with the sustained high EPU and GPR.

3 Data

Previous studies on the factors influencing oil prices demonstrate that incorporating economic and geopolitical indicators can enhance the accuracy of price predictions. Based on this, this thesis focuses on five key aspects: the benchmark price of oil, interest rates, foreign exchange rates, and indexes of economic and geopolitical uncertainty. The international derivatives oil price is benchmarked by WTI Crude Oil traded in New York, and Brent Oil traded in London (Scheitrum et al., 2018). In this thesis, only WTI Crude Oil is used as global price reference, consistently, macroeconomic factors relevant to the US market, including the exchange rate between the USD and CNY, and the economic conditions of both countries. In addition, GPR is included to be used to analyze the broader external factor, such as the impact of war on oil supply.

Data used in this thesis is from open source, including the INE, SHIBOR, Yahoo Finance, Market Watch, and Economic Policy Uncertainty index website. Since SC options opened to the market on 21 June 2021, the time span of the dataset ranges from the inception of SC options is from then to 31 December 2023.

3.1 Future and option price

Data containing the SC futures and options prices were obtained from INE [1]. The main dataset downloaded from INE is a full package of all commodities traded on the annual basis, so the primary step is the extraction of the SC-specific contracts,

[1]SC Future and Option Price, Available online: https://www.ine.cn/statements/option/thisdownload/ (accessed March 25, 2024)

then categorize them into distinct datasets for futures and options. For example, the contract name ”sc2109C455” is an option, and ”sc2109” indicates it is a future. Still working with the contract name, for options, it was deconstructed into three parts, the underlying contract (”sc2109”), type of options for call (C) or put(P), and strike price (K). In this example, it is a call option that can be exercised at the price of 455.

Another important aspect is the computation of time to maturity. Utilizing the trading guidelines listed on INE [2], the last trading day is the thirteen-to-last trading day of the month before the delivery month of the underlying SC contract, subject to the changes in national holidays. The dates of official holidays that conflicts with normal working days are also listed on INE [3]. In this case, after introducing the holidays and applying the rule to the dataset, time to maturity in days is obtained.

For the future dataset, there are two steps. The first is to calculate the historical volatility by computing the logarithm of the daily percentage changes in the 'Settle' prices of the futures data. For the use of logarithm, here it refers to Raudys and Goldstein (2022), who found that logarithmic transformations have stronger volatility predictions as measured by mean squared error and accuracy. After that, a rolling window of 30 days is applied to these log returns, and the standard deviation of the values within this window is calculated. This standard deviation is then annualized to represent an annual volatility rate. Annualization is accomplished by multiplying the standard deviation by the square root of 252, which is the typical number of trading days in a year. The second step is to match the volatility and settle price with each option's trading records. Since the option's underlying future contracts has been extracted, future contracts are used to be matched with its settle price. Based on dates and future contracts, settle price on the trading day was matched.

The WTI Crude Oil Price dataset containing the WTI Crude Oil prices was

[2]Contract Specifications, Available online: `https://www.ine.cn/eng/market/options/sc/contract/129388.html` (accessed March 28, 2024)

[3]Trading calendar, Available online: `:https://www.ine.cn/eng/search/?queryString=holiday` (accessed March 28, 2024)

sourced from Market Watch[4], and is matched with the SC option dataset by trading dates. Due to the unavailability of the settle price from the platform, the closing price is used as a substitute in the dataset to accurately reflect market sentiment for the next trading day. Although not a perfect substitute for the settle price, the closing price still provides valuable information on market conditions at the end of the trading day.

3.2 Interest rate and exchange rate

SHIBOR is the borrowing costs between prime banks in the Chinese inter-bank market and is often used as a benchmark rate for a variety of financial instruments within China, including loans, mortgages, and other derivatives. The rates include short- to long-term rates, which align closely with the underlying dynamics and pricing mechanisms of these instruments. SHIBOR's transparency and accessibility make it a reliable indicator of the money supply relationship among banks and its quotation process reflects its strong stability (Xu, 2014). Therefore, SHIBOR is adopted as a reference rate for building the option pricing models in this thesis, similar to how LIBOR has been used in financial models globally.

The data for SHIBOR [5] is available across various time scales, including 1-week, 2-week, 1-month, 3-month, 6-month, 9-month, and 1-year intervals. With the calculation of time to maturity from the previous section, SHIBOR is allocated based on the time interval, for example, if the option expires in 81 days, then it will use the rates of 3-month SHIBOR based on the trading date.

Similarly, the exchange rate data for USD to CNY was obtained from Yahoo Finance[6]. The exchange rate is key in examining the cross-market comparisons and for assessing the impact of currency fluctuations on commodity pricing.

[4] WTI Crude Oil Price, Available online: https://www.marketwatch.com/investing/future/cl.1/download-data?mod=mw_quote_tab (accessed March 25, 2024).

[5] SHIBOR, Available online: https://www.shibor.org/shibor/dataservices/ (accessed March 25, 2024).

[6] Exchange rate data, Available online: https://finance.yahoo.com/quote/CNY=X/history/ (accessed March 25, 2024).

3.3 EPU & GPR

EPU used in this thesis for China [7] and US[8] are both news-based indexes on monthly basis and are downloaded from Economic Policy Uncertainty. The methodology of this index is based on the research of Baker et al. (2016), where newspaper content is used to quantitatively capture shifts in policy-related economic uncertainty. From the website[8], the US EPU is constructed using various media sources, including analyses of reports by the Congressional Budget Office that cover tax aspects, and data from the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters, which encompasses the Consumer Price Index (CPI) and government expenditures. While the China EPU primarily sources its data from news articles that mention specific terms associated with economic uncertainty and policy decisions. From this point, US EPU is used from the news-based index.

The China EPU index uses the South China Morning Post (SCMP) with a monthly frequency count of containing specific terms related to China and economic uncertainty since 1995, and further narrowing these by including terms associated with policy discussions (Baker et al., 2013, 2016). The methodology aligns with that of the US by normalizing article frequency against the total articles published each month, standardizing this measure to an index where the average value is set to 100 for the period between January 1995 and December 2011. For validating the process, a manual check was conducted with sampling period from January 1995 to February 2012. It demonstrated that 492 out of 500 samples are correct, presenting a high accuracy with very low false positive and negative rates. This automated classification correlates strongly with manual classifications by human readers, indicating robustness in measuring policy-related economic uncertainty through media analysis.

For the US EPU, the index is broader, including 10 major newspapers, including

[7]China EPU, Available online: https://www.policyuncertainty.com/scmp_monthly.html (accessed April 3, 2024)

[8]US EPU, Available online: https://www.policyuncertainty.com/us_monthly.html (accessed April 3, 2024).

USA Today and The Wall Street Journal. The index is constructed by searching for articles that contain combinations of terms related to 'uncertainty', economic conditions, and political entities or policies. To adjust for variations in the volume of published articles over time, the count of relevant articles is normalized by the total articles published each month per paper, with these normalized values standardized to have a unit standard deviation for the period from January 1985 to December 2009.

Since EPU is monthly based, to match the daily trading data, the EPU values are matched based on the trading month but adjusted to reflect a one-month lag. For example, the trading date is June 21, 2021, then the corresponding index is May 2021.

The GPR index is downloaded from the same source as EPU[9]. Caldara and Iacoviello (2021) at the Federal Reserve Board have developed a GPR to measure the impact of adverse geopolitical events by analyzing newspaper coverage since 1900. Similar to the EPU, the index is generated through automated text searches in the archives of 10 major newspapers, counting articles related to geopolitical tensions as a proportion of total news coverage. The primary GPR index uses data starting from 1985 and fully covers the thesis period. GPR is matched using the same one-month lag data merging method.

3.4 Comparative analysis

After combining all datasets as a whole, some trends can be observed from the plots. Figure 3.1 compares the SC and WTI future prices, where WTI closing prices are presented in CNY. At first glance of the plot, the SC settle price and WTI closing prices in CNY are generally moving at the same pace. If not, then there might be arbitrage opportunity for speculators. For example, during periods of April to July 2022, when a significant price gap emerges, it presents a potential arbitrage opportunities for speculators to buy SC and sell WTI, exploiting these discrepancies with

[9]GPR, Available online: `https://www.policyuncertainty.com/gpr.html` (accessed April 3, 2024)

minimal risk management challenges. Such trends can also be quickly captured by quantitative programmatic trading systems. Additionally, the strike price (K in the figure) exhibits less fluctuation and often tracks closely with SC future, suggesting a range in which the options are frequently in-the-money. In particular, Figure 3.2 shows that although WTI price is in USD, it presents a very close pattern with the option price, indicating WTI price can be a stable reference for SC option pricing.

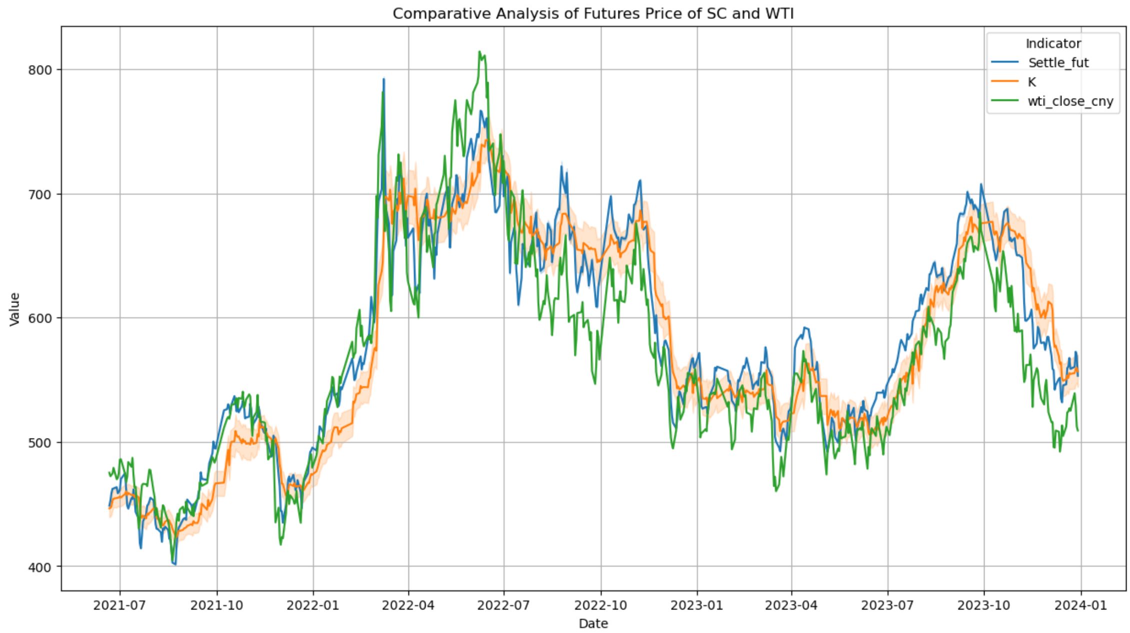

Figure 3.1: Comparative Analysis of Options and Futures of SC and WTI

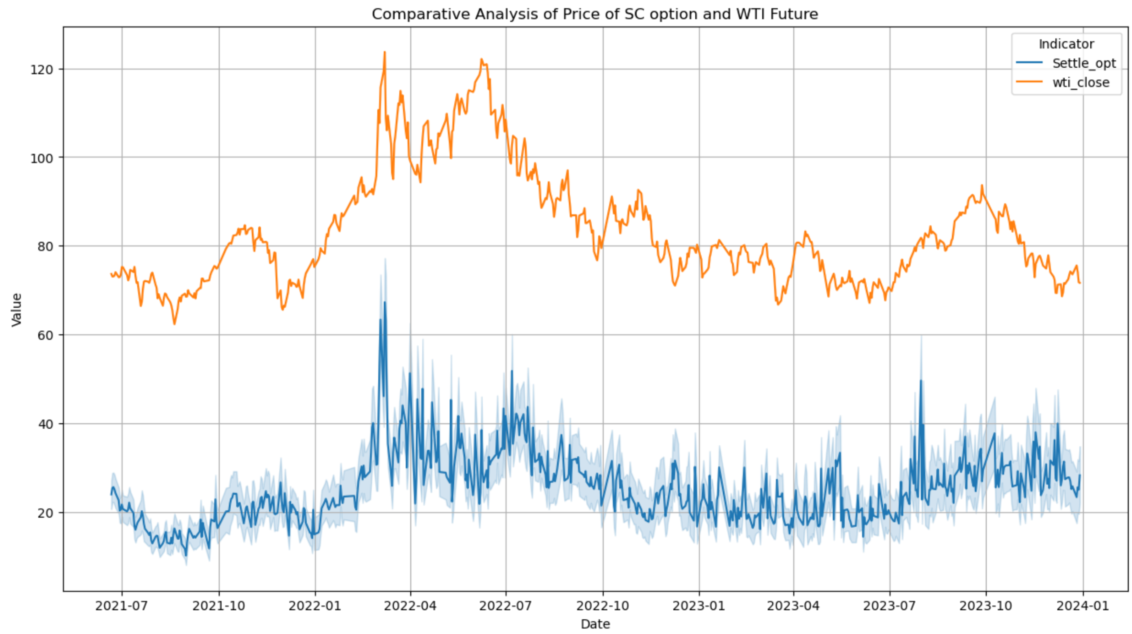

Figure 3.2: Comparative Analysis of Price of SC option and WTI Future

Figure 3.3 illustrates EPU indices for China and the US. The overall trend in both countries does not show a consistent direction over time, but rather a series of peaks and troughs that could be reactions to specific events or changes in the political-economic conditions in each country. In general, China has a more volatile policy environment as the EPU value and degree of changes are larger than the US, indicating greater sensitivity to policy-related news. In contrast, the US EPU remains relatively lower and exhibits less volatility, suggesting a more stable policy condition or less variation in policy uncertainty.

GPR measures the risk posed by geopolitical events and generally shows stable values. However, two notable spikes can be observed in April 2022 and October 2023, which may coincide with the escalation of conflicts between Russia and Ukraine, and Israel and Pakistan respectively. These spikes reflect the index's responsiveness to geopolitical tensions and their impacts on global stability.

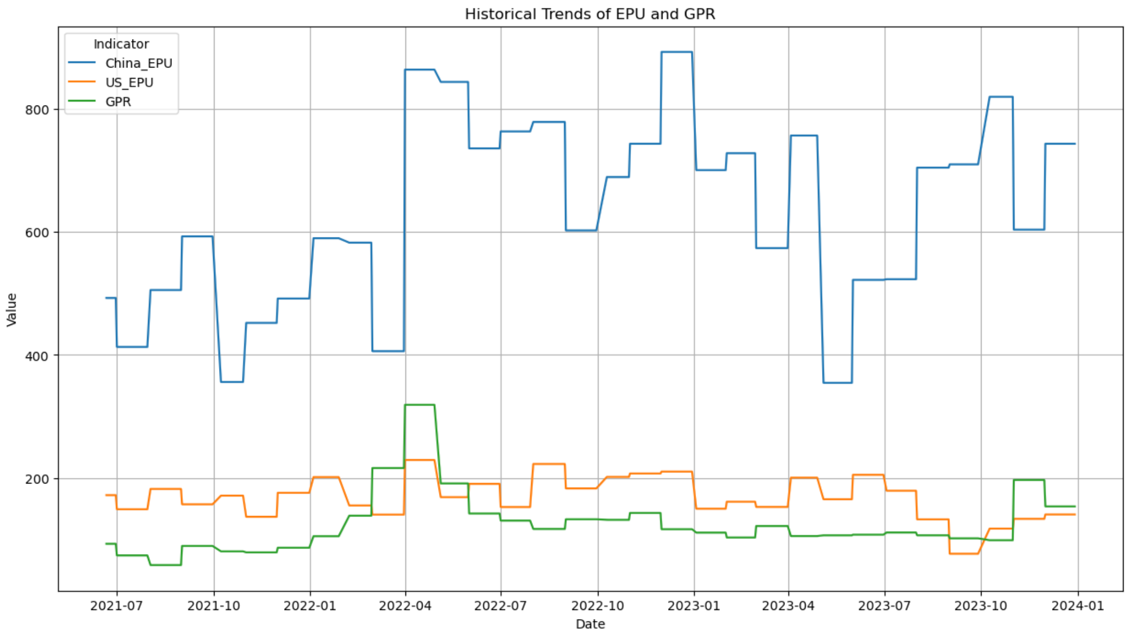

Figure 3.3: Historical Trends of EPU and GPR

4 Methodology

The methodology involves four parts to detail the processes for building models and comparing the performance. First, the stationarity test is conducted to ensure the time series data is appropriate for subsequent modeling. Then it explains the model performance metrics. The final part focuses on the models implemented, including the benchmark model and machine learning algorithms.

4.1 Stationarity test

A time series dataset is considered stationary when its statistical properties, including mean and variance, are constant over time. Testing stationarity is important for the application of many modeling techniques, as not all machine learning algorithms inherently handle time-series data. For models like RF, SVM, and XGBoostoost, stationarity is not strictly required but can help in time-series forecasting (Probst and Boulesteix, 2018; Smola and Schölkopf, 2004; Chen and Guestrin, 2016). For neural network-based models of LSTM and GRU, which can manage sequence data including non-stationary, transforming data to be stationary can enhance training efficiency and model performance (Cho et al., 2014).

To assess the stationarity of a time series, the Augmented Dickey-Fuller (ADF) test, an extension of the original Dickey-Fuller test introduced by Dickey and Fuller (1979), is employed. The ADF test extends the original approach by including higher-order regressive processes by incorporating lagged differences in the series (Said and Dickey, 1984). The null hypothesis of the ADF test is that the time series has a unit root, in other words, it is non-stationary. Conversely, the alternative hypothesis is it is stationary. For interpreting the test results, the more negative the test statistic, the smaller the p-value, and the higher the possibility of rejecting the null hypothesis, indicating the series can be considered stationary. For the critical value, there are 1%, 5%, and 10% significance levels, offering benchmarks against which the test statistic is compared. If the test statistic is more negative than the critical value, the null hypothesis is rejected at that significance level.

From Table 4.1, only strike price, time to maturity, and option settle price are stationary. To transfer the non-stationary into stationary, Box et al. (2015) state the importance of making a time series stationarity through differencing to address trends and seasonal effects, while Gujarati and Porter (2009) demonstrated log transformations are helpful for data that exhibit exponential growth, thereby stabilizing variance over time. As shown in Appendix A, Settle_fut and wti_close show clear upward or downward trends, then differencing method is applied. While Volatility and fx_close show exponential growth or non-constant variance, a log transformation followed by differencing is more proper. For the China_EPU and the US_EPU have p-values higher than 0.01 but lower than 0.05, variables are converted with difference calculation for the robustness.

Table 4.1: ADF Test Results Summary

Variable	Test Statistic	p-value	Critical Value (1%)	Stationary
K	-4.374	0.00033	-3.430	Yes
T	-14.07952	2.845e-26	-3.430	Yes
Settle_opt	-17.55898	4.106e-30	-3.430	Yes
Settle_fut	-2.130	0.232496	-3.430	No
Volatility	-2.415	0.137472	-3.430	No
fx_close	-1.168	0.687276	-3.430	No
wti_close	-2.145	0.226869	-3.430	No
China_EPU	-2.994	0.035456	-3.430	No
US_EPU	-3.190	0.020583	-3.430	No
GPR	-2.258	0.185929	-3.430	No

After transformation, the dataset satisfies the stationarity test. The results and corresponding plots are shown in Appendix A."

4.2 Model performance

To evaluate the predictive accuracy and generalization ability, Mean Absolute Error (MAE), Mean Squared Error (MSE) and R-squared (R^2) are adopted.

MAE is the average of the absolute differences between prediction and actual values. This measure is straightforward and easy to interpret, but less sensitive to the outliers due to the absolute value.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{4.1}$$

MSE is the average of the squares error, which encounter the outliers better than the MAE, as it enlarges the effect of the large differences by squaring.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{4.2}$$

R^2 is the coefficient of determination, which measures the percentage of the variance in the dependent variable that is explained by the independent variables in a model. It indicates how well the data fits to the model.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \tag{4.3}$$

4.3 Benchmark model

According to Cox et al. (1979), the Binomial Tree model assumes that at each time step, the price of the underlying asset can either go up or down, represented by the factors u and d. These values are derived from the volatility σ and time increment per step Δt. To calculate each branch value, the risk-neutral probability p is introduced to represent the likelihood that the price of the underlying asset will increase in the next time step. This probability is based on risk-neutral assumption, which assumes that all investors are risk-neutral and expect to earn at least the risk-free rate r on their investments. In this case, the expected or binomial value of the option is calculated as the weighted average of the option values at the following binomial nodes, using the probabilities p and $1-p$, then discounted back to the present using the risk-free rate. For the exercise value for a call option, the payoff is the maximum of zero or the asset price minus the strike price, and conversely for put options, which is the maximum of zero or the difference between the strike price and the asset price. Investors make decisions by comparing the binomial value and exercise value.

$$u = e^{\sigma\sqrt{\Delta t}} \tag{4.4}$$

$$d = e^{-\sigma\sqrt{\Delta t}} = \frac{1}{u} \tag{4.5}$$

$$p = \frac{e^{(r-\text{div\_yield})\Delta t} - d}{u-d} \tag{4.6}$$

$$\text{Binomial Value} = [p \times \text{Option\_up} + (1-p) \times \text{Option\_down}] \times e^{-r\Delta t} \tag{4.7}$$

where

- r - Risk-free interest rate.
- σ - Volatility of the underlying asset.
- T - Time to expiration of the option, in years.
- N - Number of time steps in the binomial model.
- div_yield - Dividend yield of the underlying asset.
- $\Delta t = \frac{T}{N}$ is the time increment per step.

This model offers simplicity and interpretability, particularly considering the condition of early exercise. Therefore, it is used as a benchmark model in this thesis. To implement the model, N is set to 100, providing a balance between sufficient iterations for accuracy and manageable computational demand. The risk-free rate is determined using SHIBOR, which aligns with market conditions, and for commodities like SC, no dividends are considered. After calculating the option value using this model, its accuracy is assessed by comparing it with actual market prices (Settle_opt) through metrics of MAE, MSE, and R^2.

4.4 Machine learning algorithms

The traditional machine learning algorithms used in this thesis include Random Forest (RF), Support Vector Machine (SVM), eXtreme Gradient Boosting (XGBoost). The deep learning algorithms adopted are Artificial Neural Network (ANN), Long Short-Term Memory (LSTM), and Gate Recurrent Unit (GRU). The adoption of these algorithms is motivated by previous research that demonstrated the effectiveness of deep learning techniques in handling complex and high-dimensional data. Tree-based models are good at dealing with non-linear relationships and can avoid overfitting problems if properly tuned. For evaluating model performance, MAE, MSE, and R^2 are adopted for all models.

RF, introduced by Breiman (2001), is an ensemble method that through a bagging technique by building multiple trees and combining them to get accurate prediction. Each tree is built with a random sample of data points, and each node is determined by a random selection of features. By averaging trees, RF can reduce the possibility of overfitting and due to the randomness, it is able to fit non-linear datasets. For building the model for this thesis dataset, hyperparameters tuning is employed. The process begins by defining a grid of hyperparameters, including combinations of the number of trees, the maximum number of features for splitting a node, the maximum depth of the trees, the minimum number of samples required to split a node, and the minimum number of samples required at each leaf node. After finding the best hyperparameters for call and put options separately, the model is trained and tested.

SVM is an algorithm primarily used for classification tasks, but it can also be effectively adapted for regression, known as Support Vector Regression (SVR). The purpose of SVR is to find a hyperplane that fits most data points within a certain distance. To handle the non-linear data, SVR employs the kernel trick including linear, polynomial, radial basis function (RBF), and sigmoid to transform the data into a higher-dimensional space where linear separation is possible. In the model, a regularization parameter (denoted by C) presents a trade-off between reducing training error and enhancing model generalization. A higher C forces the regression line to move more closely to the training data, potentially reducing the margin but risking overfitting. In this model, the hyperparameters tuning includes adjusting C and the kernel type among linear, polynomial, and RBF.

XGBoost, developed by Chen and Guestrin (2016), is an advanced implementation of gradient boosting algorithms. This method employs a series of decision trees, and each is designed to address and correct the errors made by previous ones, thereby improving prediction accuracy. To enhance model accuracy and prevent overfitting, it introduces regularization techniques including tree pruning, L1 and L2 regularization, and learning rate to avoid overfitting. Tree pruning is the process of growing the tree in full depth, and then prunes it backward to remove splits that have no additional value. L1 and L2 are helpful for adjusting the weights of the decision trees. The learning rate controls the contribution of each tree added to the model, where the smaller the value, the more gradual learning speed, and avoids fitting too closely to the training data. In this thesis, hyperparameter tuning involves searching the number of trees, the depth of the trees, and the learning rate.

ANN is a simple type of neural network architecture, which consists of input, hidden, and output layers with interconnected nodes. The input layers receive all information, and pass it to hidden layers, where a weighted sum is used through a non-linear activation function. The output layer is the one that produces predictions. For the activation function, there are sigmoid, Rectified Linear Unit (ReLU), and tanh. This structure allows ANN to capture complex patterns and relationships in data. In this thesis, the model is built into four layers in total with dropout layers. For the model tuning, the number of units, dropout rates, and learning rate are examined. In the training process, early stopping is added for mitigating overfitting problems.

LSTM, developed by Hochreiter and Schmidhuber (1997), is a type of RNN that is good at handling sequence predictions. While standard RNN is able to handle the sequence prediction, but not for long sequences due to the vanishing gradient, as it is hard to retain information from earlier time steps. LSTM solves this problem by including gates that can regulate the flow of information by retaining or discarding the information. The gate includes input, forget and output, in this case, LSTM can preserve long-term dependencies and mitigate the vanishing gradient problem. The basic model architecture is the same as ANN of four layers, while hyperparameter tuning involves selecting the number of units in each layer, and the dropout rate to prevent overfitting.

Similar to LSTM, GRU is another type of RNN that aims to solve the vanishing gradient problem by adopting gating mechanisms to control and manage the flow of information between cells in the network, but with fewer parameters. GRU combines the input and forget gates into a single gate, making it simpler and often faster to train than LSTM, particularly on smaller datasets. The model is tuned and evaluated using the same method as LSTM to ensure consistency and comparability.

The hyperparameter tuning for the models is summarized in the tables below:

Model	Hyperparameters	Description	Range
RF	n_estimators	Number of trees in the forest.	50, 100, 200, 500
	max_features	Number of features to consider when looking for the best split.	'auto', 'sqrt'
	max_depth	Maximum depth of the tree.	10, 20, 30, 50
	min_samples_split	Minimum number of samples required to split an internal node.	2, 5, 10, 20
	min_samples_leaf	Minimum number of samples required to be at a leaf node.	1, 2, 4, 10
SVR	C	Regularization parameter. The strength of the regularization is inversely proportional to C.	0.1, 1, 10
	kernel	Specifies the kernel type to be used in the algorithm.	'rbf', 'linear', 'poly'
XGBoost	n_estimators	Number of gradient boosted trees.	50, 100, 200, 500
	learning_rate	Step size shrinkage used in update to prevent overfitting.	0.01, 0.05, 0.1, 0.5
	max_depth	Maximum depth of a tree. Increasing this value will make the model more complex and more likely to overfit.	10, 20, 30, 50

Table 4.2: Hyperparameters for RF, SVR, and XGBoost models.

Model	**Layer 1**	**Layer 2**	**Layer 3**	**Output Layer**	**Learning Rate**
ANN	Units 30-100 (step=10)	Dropout 0-0.5 (step=0.1)	Units 30-100 (step=10)	1	1e-4 to 1e-2
LSTM	Units 30-100 (step=10)	Dropout 0-0.5 (step=0.1)	Units 30-100 (step=10)	1	1e-4 to 1e-2
GRU	Units 30-100 (step=10)	Dropout 0-0.5 (step=0.1)	Units 30-100 (step=10)	1	1e-4 to 1e-2

Table 4.3: Hyperparameter tuning for ANN, LSTM, and GRU models.

5 Results and Discussions

5.1 Hyperparameters tuning

The optimized hyperparameters after tuning are summarized in Table 5.1. For the RF model, the max depth for Call options is greater than for Put options, which may indicate a more complex pattern within the Call options data. In contrast, the SVR parameters are the same for both Call and Put options, suggesting a similar complexity in the decision boundaries. The use of the RBF kernel in both cases helps capture non-linear relationships effectively. For XGBoost, both models maintain consistent parameters: a learning rate of 0.1, 500 estimators, and a maximum depth of 10. This uniformity suggests a balanced approach to managing model complexity and training efficiency across both types of options. In particular, the number of estimators and units in all models tend to reach the upper limits set within the tuning process, which could potentially influence the final performance. However, due to computational efficiency constraints, adjustments to these tuning parameters were not further explored.

5.2 Performance comparasion

In the case of LSTM, there is a slight difference in the number of units between Call (80 units) and Put (100 units) options, indicating a need for additional capacity to handle more complex temporal patterns in Put options. The GRU models have 100 units for both option types, suggesting that this level of complexity is sufficient.

While the LSTM models present different learning rates, the GRU and ANN

Table 5.1: Parameters for Machine Learning Models

Model	Parameters
RF Call	max_depth : 50, max_features : sqrt , min_samples_leaf : 1, min_samples_split : 2, n_estimators : 500
RF Put	max_depth : 30, max_features : sqrt , min_samples_leaf : 1, min_samples_split : 2, n_estimators : 500
SVR Call	C : 10, kernel : rbf
SVR Put	C : 10, kernel : rbf
XGBoost Call	learning_rate : 0.1, max_depth : 10, n_estimators : 500
XGBoost Put	learning_rate : 0.1, max_depth : 10, n_estimators : 500
ANN Call	units : 80, dropout : 0.0, learning_rate : 0.000338
ANN Put	units : 80, dropout : 0.0, learning_rate : 0.000338
LSTM Call	units : 80, dropout : 0.0, learning_rate : 0.0056
LSTM Put	units : 100, dropout : 0.0, learning_rate : 0.0097
GRU Call	units : 100, dropout : 0.0, learning_rate : 0.0057
GRU Put	units : 100, dropout : 0.0, learning_rate : 0.0057

models keep the same, indicating the learning processes are varied. Despite this, all deep learning models show a dropout rate of zero, implying no regularization through dropout. In contrast to the deep learning models, RF and XGBoost manage model complexity through adjustments in tree depth and the number of trees, directly influencing how the models handle overfitting.

For the model performance of Call options, XGBoost presents the lowest errors and highest R^2 score among all models, showing superior predictive accuracy than the benchmark Binomial Tree model. RF presents robustness with moderate error metrics and a solid R^2 score, and is slightly better than the benchmark. ANN has better performance than LSTM and GRU but still does not reach the effectiveness of RF or XGBoost. SVR exhibits the poorest performance with significantly high error and low R^2 score.

Table 5.2: Performance Metrics of the Benchmark Binomial Tree Model

Metric	Call Options	Put Options
Mean Absolute Error	10.1891	9.1592
Mean Squared Error	218.779	194.130
R^2 Score	0.7807	0.8093

Table 5.3: Performance of Machine Learning Models for Call Options

Model	MAE	MSE	R^2
RF	8.093491	207.292088	0.795333
SVR	16.882048	786.933709	0.223031
XGBoost	4.023227	90.589526	0.910558
ANN	13.479388	342.052246	0.662279
LSTM	11.815427	398.892242	0.606159
GRU	12.670074	434.812042	0.570694

For the model performance of Put options, all models present stronger performances with lower errors and a higher R^2 score than Call options, implying it is more effective at capturing the patterns of Put options. The top performance model maintains to be XGBoost that exceeds other models and outperforms the benchmark. Within the deep learning group, this time is LSTM that has a better metrics than ANN and GRU, implying a stronger predictive ability and accuracy.

Table 5.4: Performance of Machine Learning Models for Put Options

Model	MAE	MSE	R^2
RF	6.944702	155.179787	0.855177
SVR	11.781617	395.767961	0.630646
XGBoost	3.179414	57.824661	0.946035
ANN	11.914616	290.482483	0.728905
LSTM	8.497591	205.748596	0.807983
GRU	9.858852	251.656342	0.765140

For the model predictive ability, the results can also be interpreted visually (Appendix B). Time-series plots for RF, SVR and XGBoost display the predicted and actual prices over time. Consistent with the performance metrics, XGBoost presents a high overlap between the predicted and actual points including the outliers. RF also shows a rather robust performance but this is not the case for SVR, where it shows significant discrepancies. For the deep learning algorithms, LSTM and GRU present close results for the capturing main trends but are not consistent at value extremes. The ANN model, in this case, has a downward tail away from the accuracy line. This divergence likely contributes to the observed higher error rates and lower R^2 score, indicating specific areas where the model's predictions are less accurate.

For deep learning algorithms training and validation loss trends (Appendix B), the maximum epochs are set to 50, while with early stopping, all models stopped under 10 epochs. Generally, LSTM models show more stable learning curves with less variability in validation loss, especially for Call options. In contrast, ANN and GRU models exhibit more dramatic fluctuations in validation loss, indicating issues with overfitting at various points during training. In addition, the spikes in validation loss for the ANN and LSTM models on Put options might suggest that these models are more sensitive to the specific characteristics or noise within the training data for Put options.

Overall, XGBoost and RF are the most robust and high-performing models. XGBoost consistently outperforms the benchmark across both Call and Put options, suggesting its strong adaptability and accuracy in financial modeling. RF also generally performs comparably to or better than the benchmark, especially in Put options. Deep learning models tend to underperform against the benchmark in Call options but show competitive or better performance in Put options. SVR, despite its generally poor performance, improves in Put options but remains below the benchmark and other models.

6 Conclusion

This thesis explores the application of machine learning techniques to SC options pricing by involving traditional model features and macroeconomic factors. By constructing and analyzing a Binomial Tree model for American options as the benchmark model and comparing the outcomes from machine learning algorithms, this thesis tried to evaluate the potential improvement in option pricing using macroeconomic factors.

The machine learning algorithms employed include tree-based and deep learning techniques. While models share some similarities, the results varied from the performance metrics. Among all, XGBoost outperformed the benchmark Binomial Tree model across both Call and Put options. This superior performance highlights XGBoost's robustness and its ability to capture complex, non-linear relationships that traditional models might miss. However, other models did not consistently outperform the benchmark, leading to a question regarding the direct contribution of macroeconomic factors to enhance the predictive accuracy of pricing models. This suggests that while the macroeconomic factors can enrich model inputs to some extent, it is unclear how effectively specific algorithms can utilize this information.

In conclusion, this thesis provides insights into how macroeconomic factors influence option pricing, demonstrating that XGBoost significantly outperforms the benchmark model. Although not all models showed significant performance results, future improvements could involve adopting alternative data sources or expanding the range of hyperparameters tuned to further enhance model performance.

Bibliography

Aloui, R., Gupta, R. and Miller, S. M. (2016), 'Uncertainty and crude oil returns', *Energy Economics* **55**, 92–100.

Amano, R. A. and Van Norden, S. (1998), 'Exchange rates and oil prices', *Review of international economics* **6**(4), 683–694.

Antonakakis, N., Gupta, R., Kollias, C. and Papadamou, S. (2017), 'Geopolitical risks and the oil-stock nexus over 1899–2016', *Finance Research Letters* **23**, 165–173.

Baker, S., Bloom, N. and Davis, S. J. (2016), 'Measuring economic policy uncertainty', *Quarterly Journal of Economics* **131**(4).

Baker, S., Bloom, N., Davis, S. J. and Wang, X. (2013), Economic policy uncertainty in china. Unpublished paper, University of Chicago.

Balcilar, M., Bekiros, S. and Gupta, R. (2017), 'The role of news-based uncertainty indices in predicting oil markets: a hybrid nonparametric quantile causality method', *Empirical Economics* **53**, 879–889.

Basher, S. A., Haug, A. A. and Sadorsky, P. (2012), 'Oil prices, exchange rates and emerging stock markets', *Energy Economics* **34**(1), 227–240.

Bennell, J. and Sutcliffe, C. (2004), 'Black–scholes versus artificial neural networks in pricing ftse 100 options', *Intelligent Systems in Accounting, Finance and Management* **12**(4), 243–260.

Black, F. and Scholes, M. (1973), 'The pricing of options and corporate liabilities', *Journal of Political Economy* **81**, 637–659.

Box, G. E., Jenkins, G. M., Reinsel, G. C. and Ljung, G. M. (2015), *Time series analysis: forecasting and control*, John Wiley & Sons.

Boyle, P. (1986), 'Options: A monte carlo approach', *Journal of Financial Economics* **4**(3), 323–338.

Boyle, P. P. (1977), 'Options: A monte carlo approach', *Journal of Financial Economics* **4**(3), 323–338.

Breiman, L. (2001), 'Random forests', *Machine learning* **45**, 5–32.

Brennan, M. J. and Schwartz, E. S. (1977), 'The valuation of american put options', *The Journal of Finance* **32**(2), 449–462.

Buehler, H., Gonon, L., Teichmann, J. and Wood, B. (2019), 'Deep hedging', *Quantitative Finance* **19**(8), 1271–1291.

Caldara, D. and Iacoviello, M. (2021), 'Measuring geopolitical risk', *Working Paper, Board of Governors of the Federal Reserve Board* . Working paper.

Chen, T. and Guestrin, C. (2016), Xgboost: A scalable tree boosting system, *in* 'Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining', pp. 785–794.

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. and Bengio, Y. (2014), 'Learning phrase representations using rnn encoder-decoder for statistical machine translation', *arXiv preprint arXiv:1406.1078* .

Colgan, J. D. (2014), 'The emperor has no clothes: The limits of opec in the global oil market', *International Organization* **68**(3), 599–632.

Cox, J. C., Ross, S. A. and Rubinstein, M. (1979), 'Option pricing: A simplified approach', *Journal of Financial Economics* **7**(3), 229–263.

Culkin, R. and Das, S. (2017), 'Machine learning in finance: the case of deep learning for option pricing', *Working Paper, Santa Clara University* .

Cunado, J., Gupta, R., Lau, C. K. M. and Sheng, X. (2020), 'Time-varying impact of geopolitical risks on oil prices', *Defence and Peace Economics* **31**(6), 692–706.

Dickey, D. A. and Fuller, W. A. (1979), ‘Distribution of the estimators for autoregressive time series with a unit root’, *Journal of the American statistical association* **74**(366a), 427–431.

Engle, R. F. (1982), ‘Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation’, *Econometrica: Journal of the econometric society* pp. 987–1007.

Fattouh, B. and Sen, A. (2016), ‘The past, present, and future role of opec’, *The Palgrave handbook of the international political economy of energy* pp. 73–94.

Gençay, R. and Salih, A. (2003), ‘Degree of mispricing with the black-scholes model and nonparametric cures’, *Annals of Economics and Finance* pp. 73–101.

Golub, S. S. (1983), ‘Oil prices and exchange rates’, *The Economic Journal* **93**(371), 576–593.

Gujarati, D. N. and Porter, D. C. (2009), *Basic econometrics*, McGraw-hill.

Hamilton, J. D. (2003), ‘What is an oil shock?’, *Journal of econometrics* **113**(2), 363–398.

Heston, S. L. (1993), ‘A closed-form solution for options with stochastic volatility with applications to bond and currency options’, *The review of financial studies* **6**(2), 327–343.

Hochreiter, S. and Schmidhuber, J. (1997), ‘Long short-term memory’, *Neural computation* **9**(8), 1735–1780.

Hu, M., Zhang, D., Ji, Q. and Wei, L. (2020), ‘Macro factors and the realized volatility of commodities: a dynamic network analysis’, *Resources Policy* **68**, 101813.

Hull, J. and White, A. (1987), ‘The pricing of options on assets with stochastic volatilities’, *The journal of finance* **42**(2), 281–300.

Hutchinson, J. M., Lo, A. W. and Poggio, T. (1994), ‘A nonparametric approach to pricing and hedging derivative securities via learning networks’, *Journal of Finance* **49**(3), 851–889.

Jo, S. (2012), The effects of oil price uncertainty on the macroeconomy, Technical report, Bank of Canada working paper.

Kilian, L. (2009), 'Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market', *American Economic Review* **99**(3), 1053–1069.

Li, W. (2022), Application of machine learning in option pricing: A review, *in* '2022 7th International Conference on Social Sciences and Economic Development (ICSSED 2022)', Atlantis Press, pp. 209–214.

Liu, Y. and Zhang, X. (2023), 'Option pricing using lstm: A perspective of realized skewness', *Mathematics* **11**(2), 314.
URL: *https://doi.org/10.3390/math11020314*

Longstaff, F. A. and Schwartz, E. S. (2001), 'Valuing american options by simulation: a simple least-squares approach', *The review of financial studies* **14**(1), 113–147.

Merton, R. C. (1973), 'Theory of rational option pricing', *Bell Journal of Economics and Management Science* **4**, 141–183.

Probst, P. and Boulesteix, A.-L. (2018), 'To tune or not to tune the number of trees in random forest', *Journal of Machine Learning Research* **18**(181), 1–18.

Raudys, A. and Goldstein, E. (2022), 'Forecasting detrended volatility risk and financial price series using lstm neural networks and xgboost regressor', *Journal of Risk and Financial Management* **15**(12), 602.

Said, S. E. and Dickey, D. A. (1984), 'Testing for unit roots in autoregressive-moving average models of unknown order', *Biometrika* **71**(3), 599–607.

Scheitrum, D. P., Carter, C. A. and Revoredo-Giha, C. (2018), 'Wti and brent futures pricing structure', *Energy Economics* **72**, 462–469.

Shanghai International Energy Exchange (INE) (2024), '2024 Development Report of Shanghai Crude Oil Futures and Options Market', `https://www.ine.cn/upload/20240325/1711336456532.pdf`. [Accessed 15 April 2024].

Smola, A. J. and Schölkopf, B. (2004), 'A tutorial on support vector regression', *Statistics and computing* **14**, 199–222.

Wei, Y., Liu, J., Lai, X. and Hu, Y. (2017), 'Which determinant is the most informative in forecasting crude oil market volatility: Fundamental, speculation, or uncertainty?', *Energy Economics* **68**, 141–150.

Xu, Z. H. (2014), 'The analysis of the effectiveness of the shibor as a market benchmark interest rate', *Advanced Materials Research* **850**, 1016–1019.

Yao, J., Li, Y. and Tan, C. L. (2000), 'Option price forecasting using neural networks', *Journal of Forecasting* **17**(5-6), 369–388.

Yi, A., Yang, M. and Li, Y. (2021), 'Macroeconomic uncertainty and crude oil futures volatility–evidence from china crude oil futures market', *Frontiers in Environmental Science* **9**, 636903.

Zhang, J. and Huang, W. (2021), 'Option hedging using lstm-rnn: an empirical analysis', *Quantitative Finance* **21**(10), 1753–1772.

A ADF Test Results

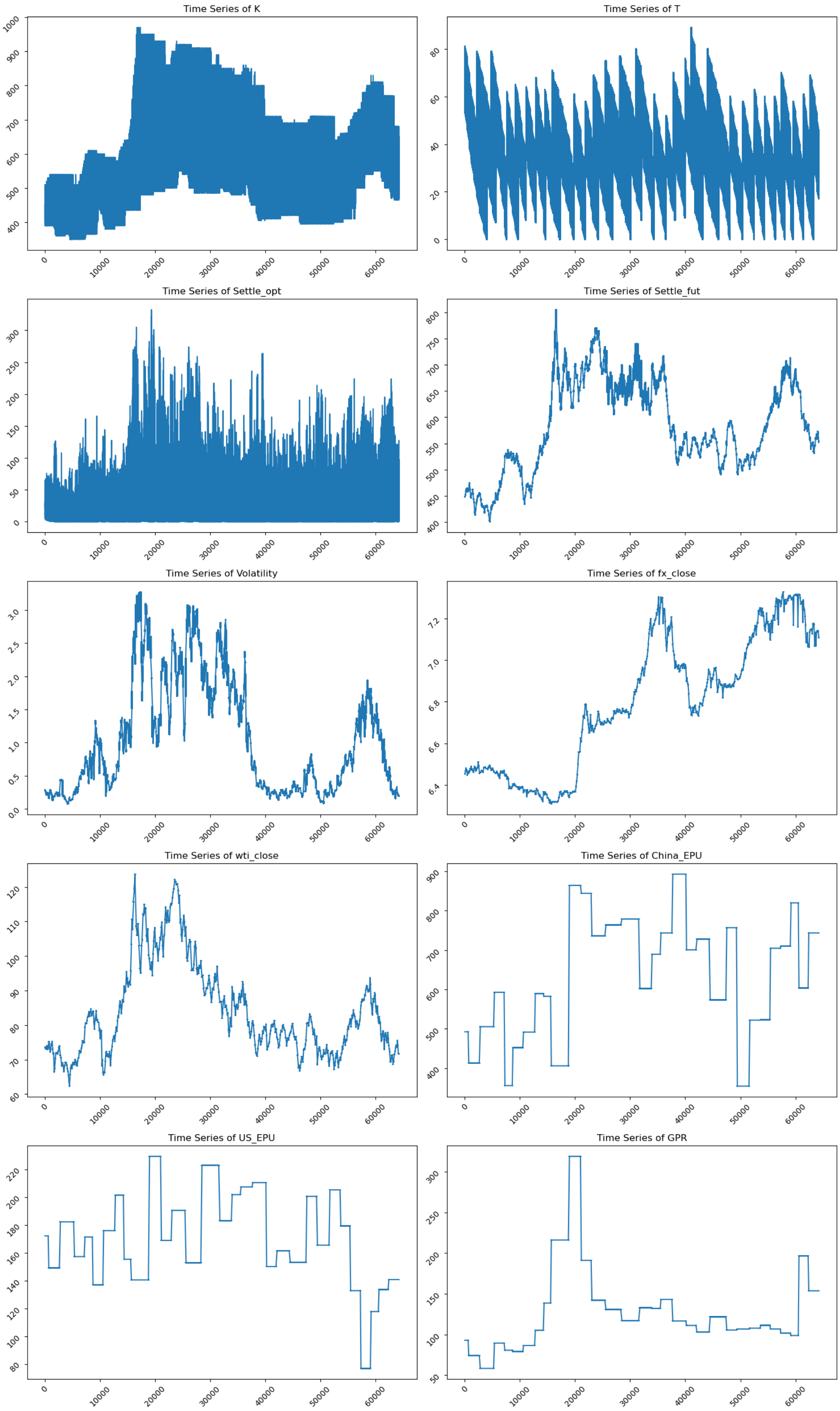

Figure A.1: *Time Series Plot of Features*

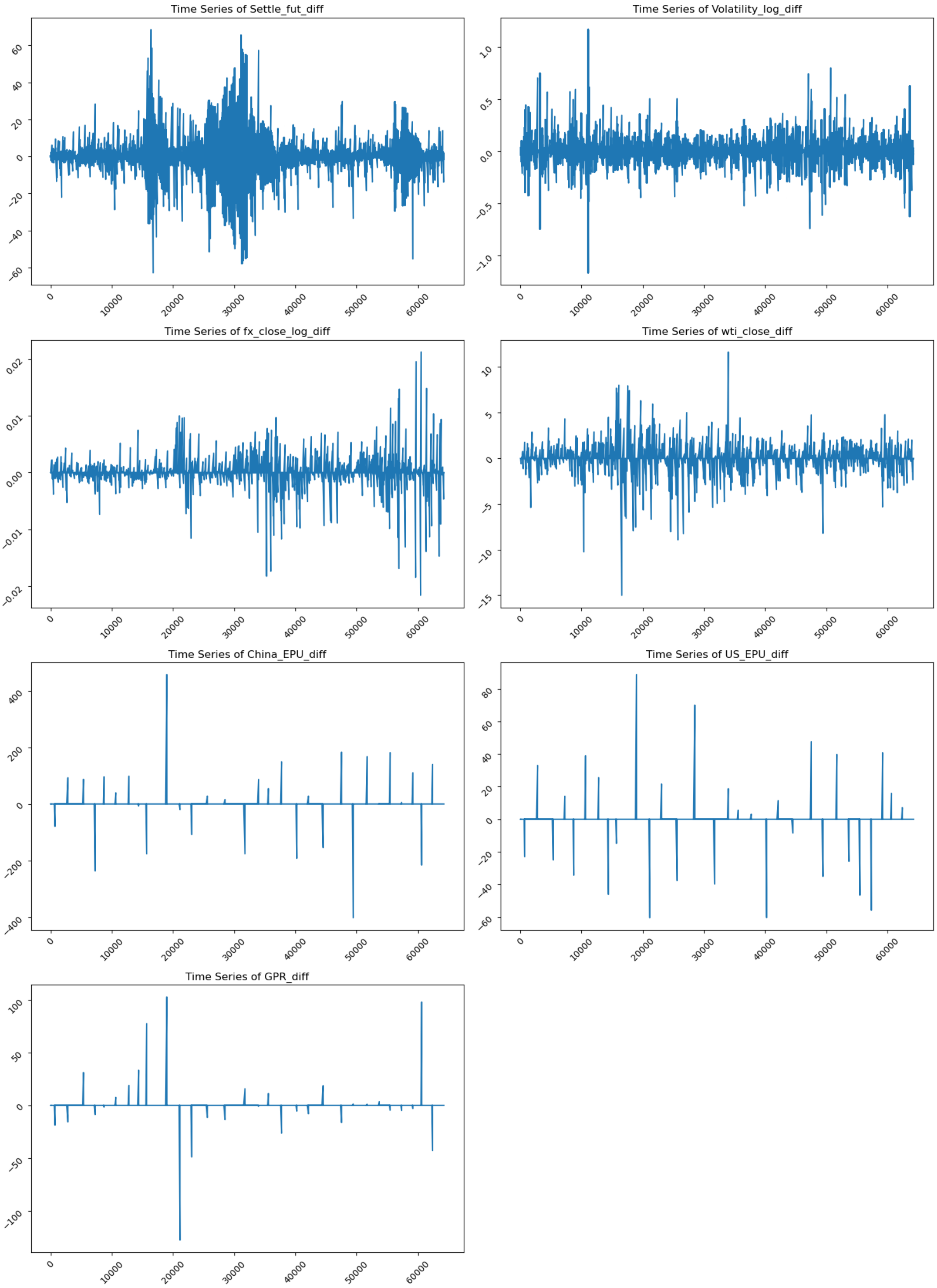

Figure A.2: *Time Series Plot of Transformed Features*

Table A.1: ADF Test Results for Differentiated and Log-Differentiated Data

Variable	Test Statistic	p-value	Critical Value (1%)	Stationary
Settle_fut_diff	-34.944491	0.000000	-3.430	Yes
Volatility_log_diff	-33.480714	0.000000	-3.430	Yes
fx_close_log_diff	-253.475411	0.000000	-3.430	Yes
wti_close_diff	-253.471897	0.000000	-3.430	Yes
China_EPU_diff	-253.472197	0.000000	-3.430	Yes
US_EPU_diff	-253.471978	0.000000	-3.430	Yes
GPR_diff	-253.472172	0.000000	-3.430	Yes

B Model Performances Plots

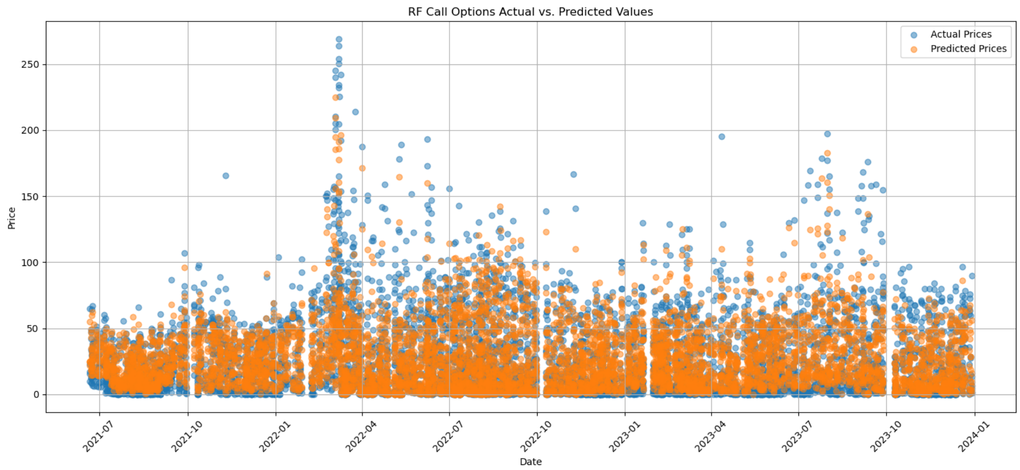

Figure B.1: *RF Call Options Actual vs. Predicted Values*

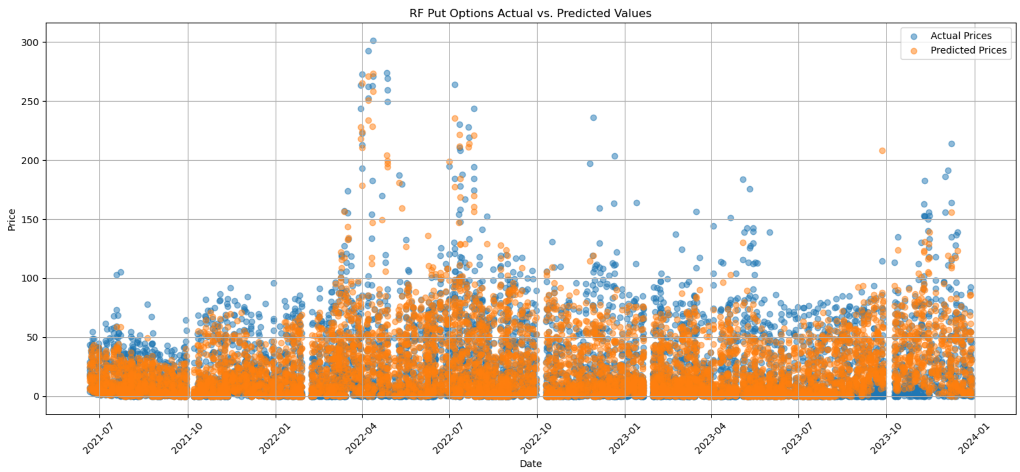

Figure B.2: *RF Put Options Actual vs. Predicted Values*

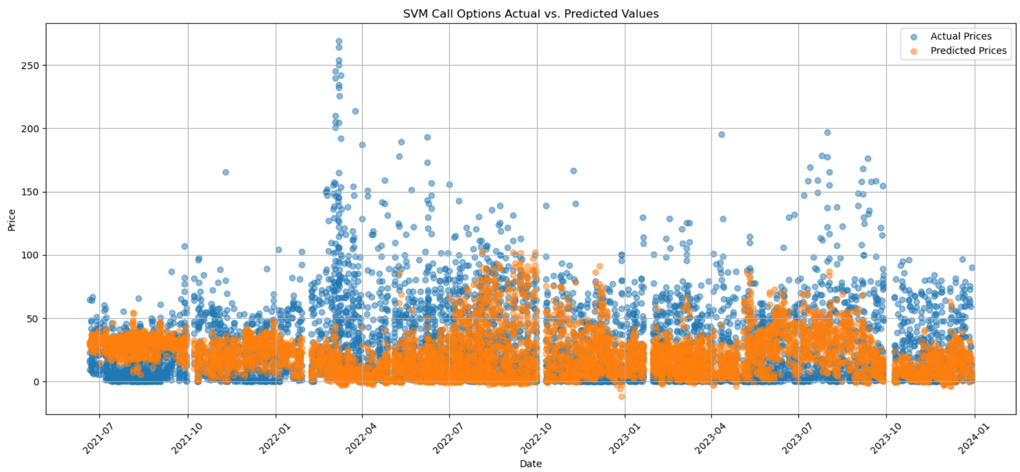

Figure B.3: *SVM Call Options Actual vs. Predicted Values*

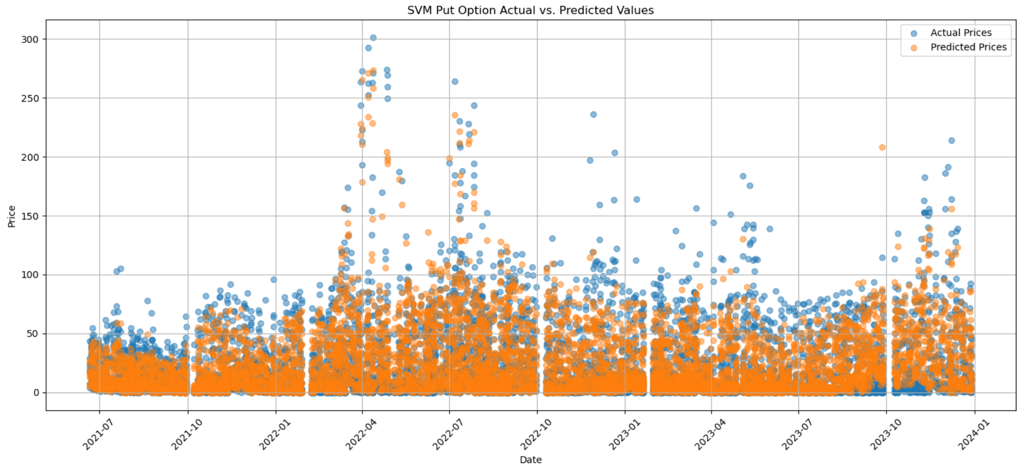

Figure B.4: *SVM Put Options Actual vs. Predicted Values*

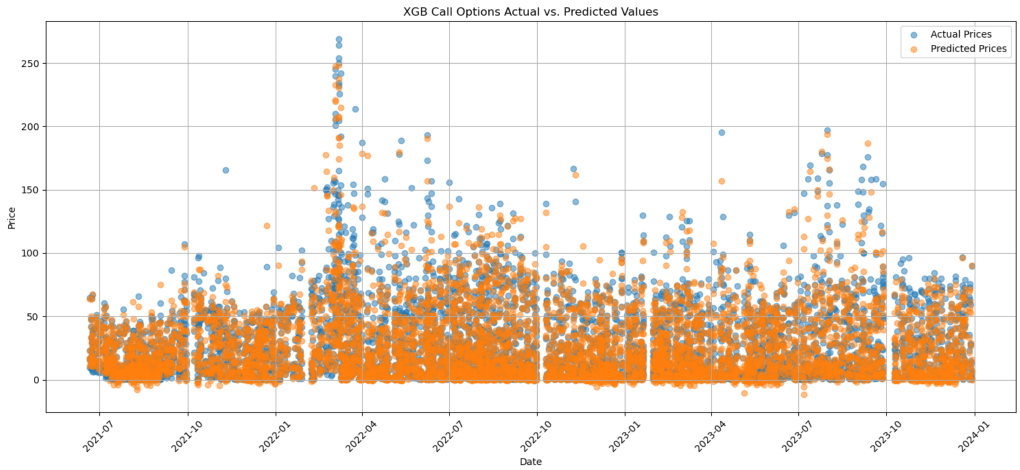

Figure B.5: *XGBoost Call Options Actual vs. Predicted Values*

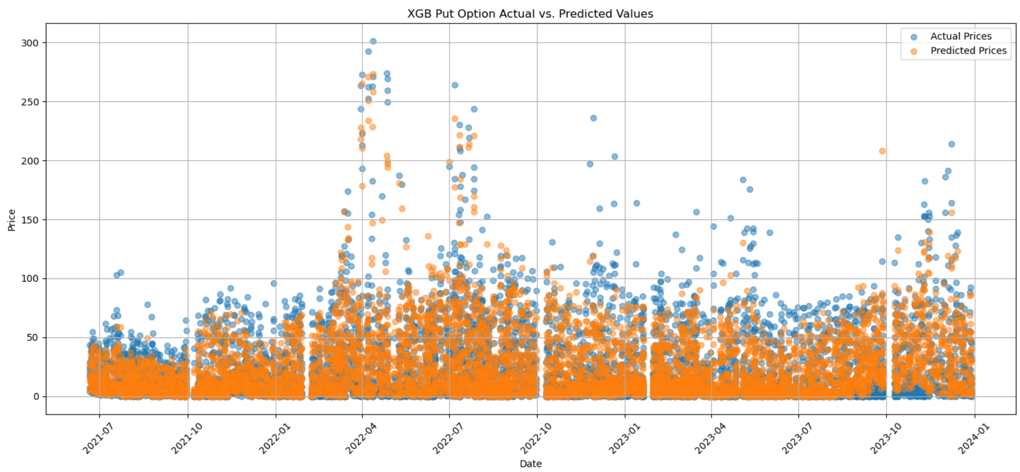

Figure B.6: *XGBoost Put Options Actual vs. Predicted Values*

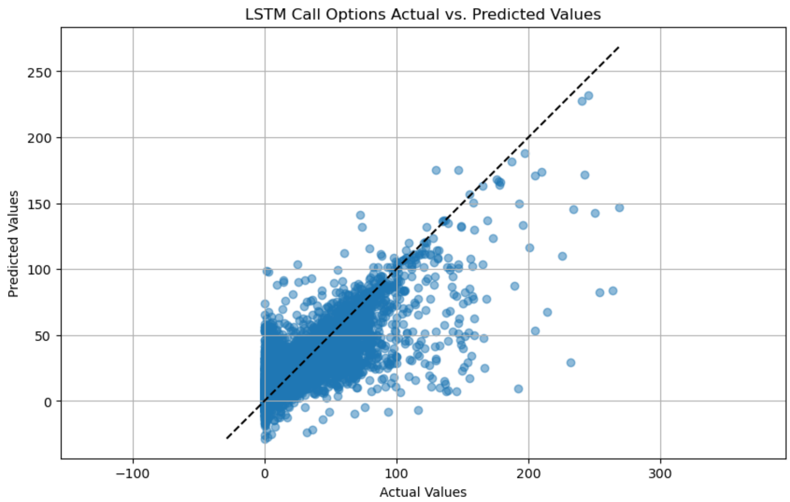

Figure B.7: *LSTM Call Options Actual vs. Predicted Values*

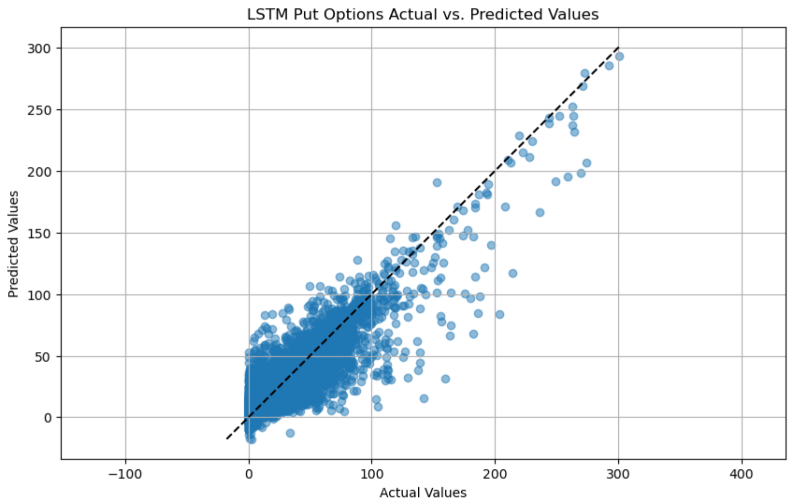

Figure B.8: *LSTM Put Options Actual vs. Predicted Values*

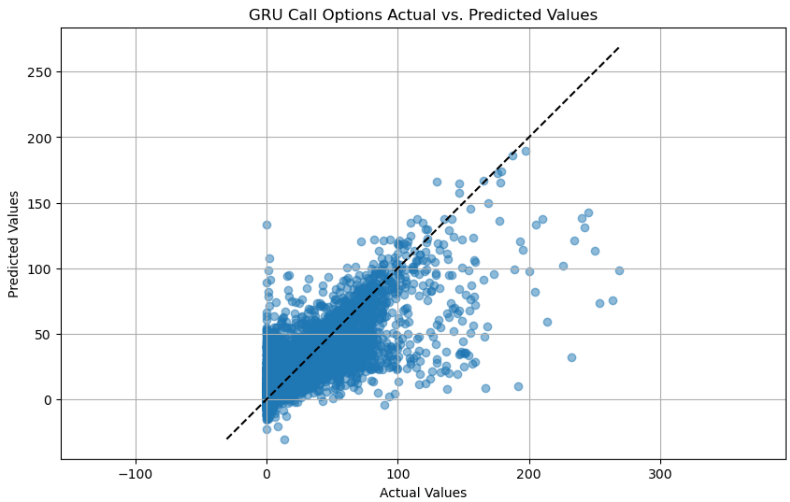

Figure B.9: *GRU Call Options Actual vs. Predicted Values*

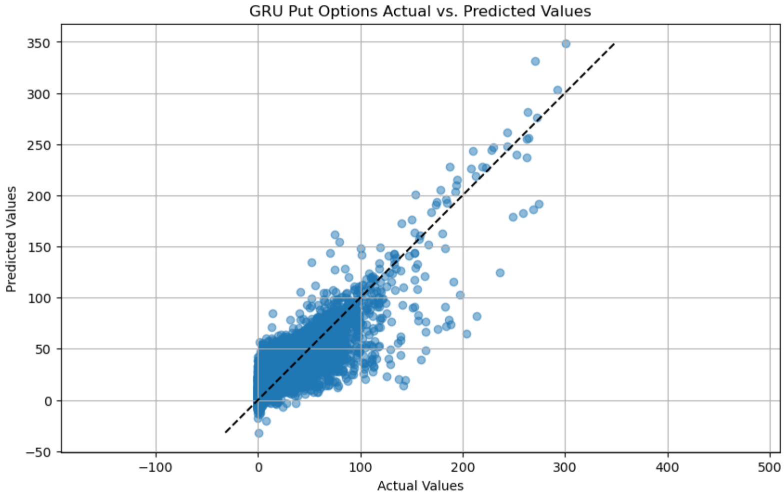

Figure B.10: *GRU Put Options Actual vs. Predicted Values*

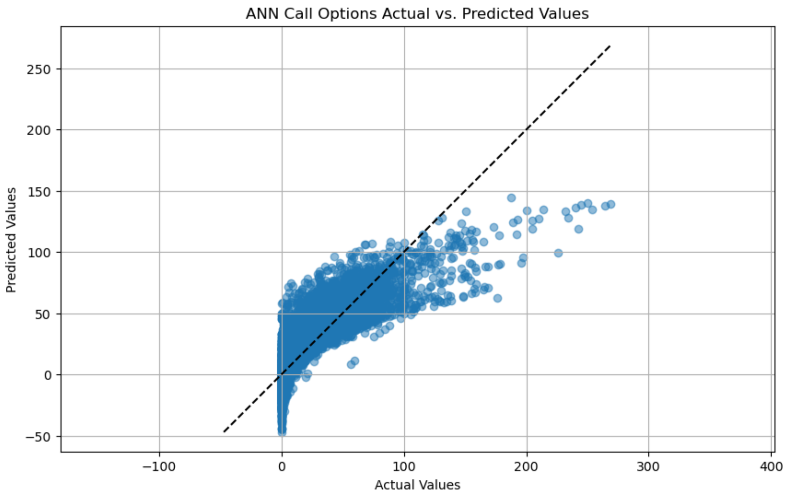

Figure B.11: *ANN Call Options Actual vs. Predicted Values*

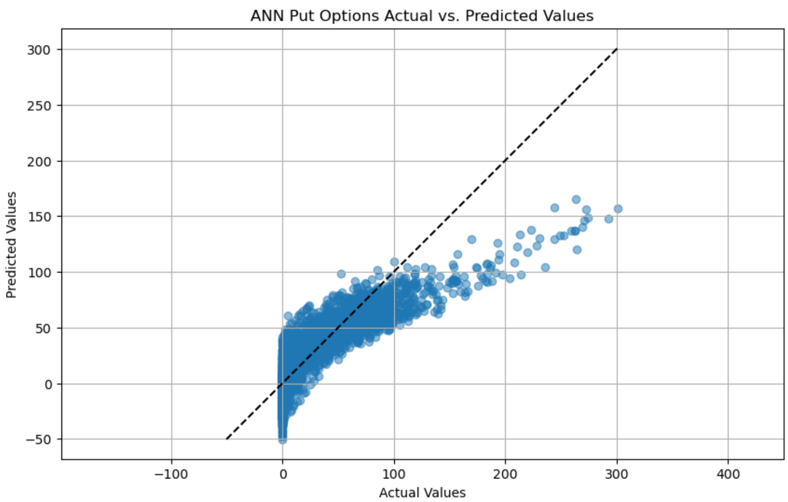

Figure B.12: *ANN Put Options Actual vs. Predicted Values*

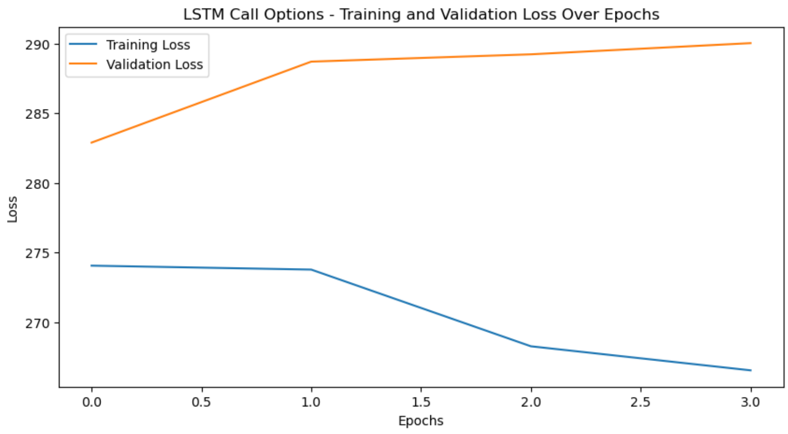

Figure B.13: *LSTM Call Options - Training and Validation Loss Over Epochs*

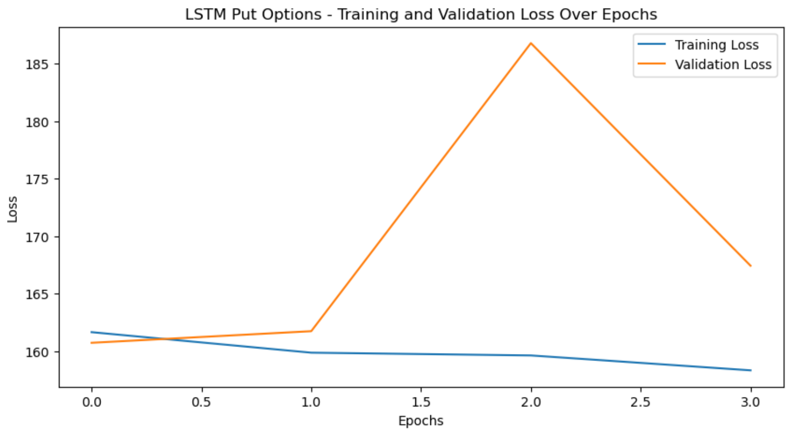

Figure B.14: *LSTM Put Options - Training and Validation Loss Over Epochs*

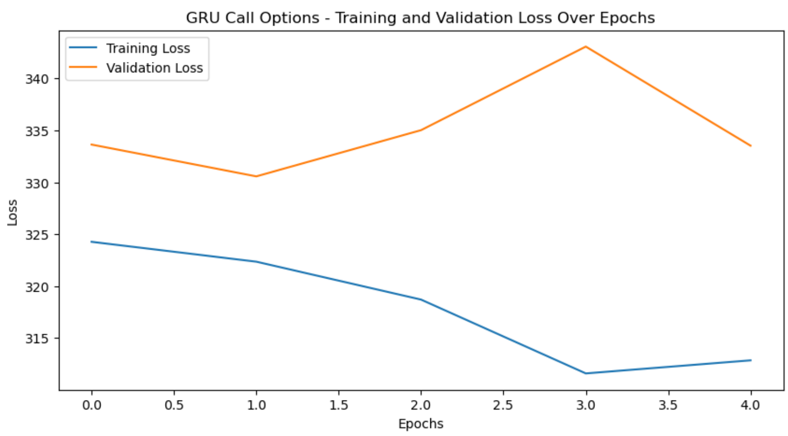

Figure B.15: *GRU Call Options - Training and Validation Loss Over Epochs*

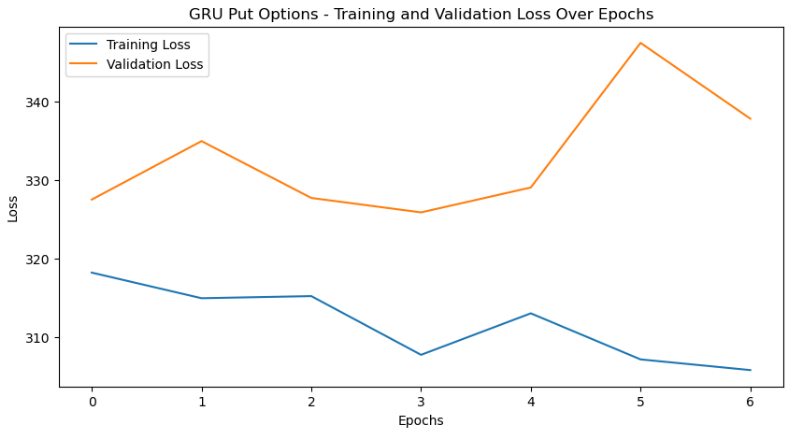

Figure B.16: *GRU Put Options - Training and Validation Loss Over Epochs*

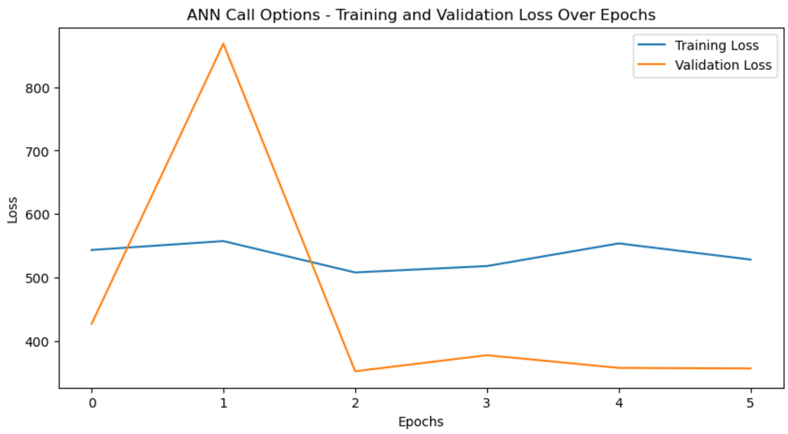

Figure B.17: *ANN Call Options - Training and Validation Loss Over Epochs*

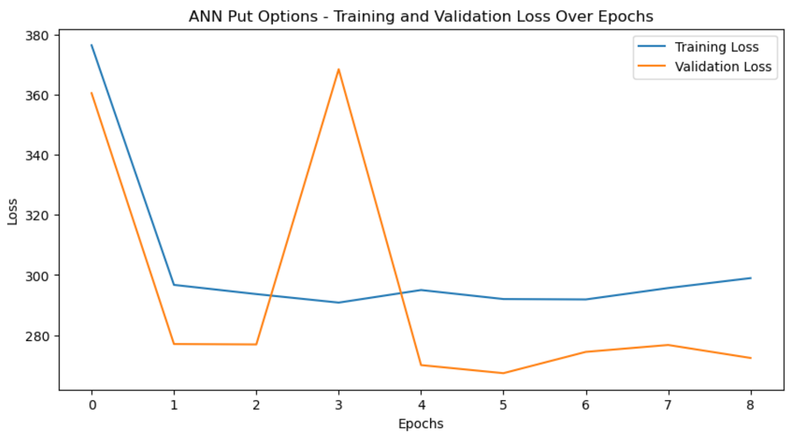

Figure B.18: *ANN Put Options - Training and Validation Loss Over Epochs*