ERASMUS SCHOOL OF ECONOMICS

MASTER THESIS QUANTITATIVE FINANCE

Machine Learning-Based Modeling of the Implied Volatility Surface

Name student: Anastasiia Borisova Student ID: 620860

Supervisor: Dr. Gustavo Freire Second assessor: Dr. Maria Grith

April 30, 2023

Abstract

Precise forecasts of implied volatility provide numerous opportunities to enhance financial decisions and trading systems. This study investigates a two-step approximation approach to predict the implied volatility surface. We utilize BS and Heston models as our parametric implied volatility models. We enhance their performance with neural networks, namely FNN, LSTM, GRU, and CNN. Our findings reveal that the BS correction provides better estimations of implied volatility estimates than the Heston model in terms of RSME. Additionally, time-dependent neural networks outperformed their benchmarks, particularly FNN. LSTM produced more stable results, while CNN and GRU displayed unstable performances. We further expand our research by incorporating time-variant covariates, such as VIX, LTP and LTV, realized volatility, EPU, ADS, term spread, and credit spread, in the neural networks. We observed that macro-related features worsened the models' predictions, while moneyness, time to maturity, VIX, LTP, and LTV significantly improved performance.

Contents

List of Tables

List of Figures

1 Introduction

Implied volatility (IV) is an indispensable tool for traders and market makers worldwide, representing the market's anticipation of the underlying asset's future movement. In other words, it reflects the level of uncertainty or risk that the market associates with an asset's future price direction. By definition, IV is the volatility level σ that equates the Black-Scholes option price to its market price S, for an option with a strike K and a period to maturity τ. It is critical to distinguish implied volatility from historical volatility, which is a measure of an underlying asset's past activity. In contrast, IV reflects the market's perception of the asset's future price direction. For options that are at-the-money (ATM), a high IV value suggests the expectation of significant anticipated changes in the underlying asset's price in the future. Conversely, low IV values indicate the market expects little or no significant price changes. Additionally, the implied volatility surface (IVS), a function of the implied volatility of the strike price K and time to maturity τ, is a key representation of the market's expectation of an underlying asset's future price movements. In practice, the implied volatility surface violates the Black-Scholes model assumptions as options that are out-of-the-money (OTM) have higher implied volatility than at-the-money options, creating the so-called "smile." Despite the existence of stochastic volatility models, such as the Heston model, which accounts for the volatility of the underlying asset based on a random process, pricing errors can still affect even parametric models that suggest stochastic volatility behavior. Therefore, the field of IVS prediction requires further research to improve the precision and robustness of models.

Numerous IV evaluation models have been proposed in recent years, but there is a growing demand for more accurate and reliable approaches. In this paper, we examine an innovative two-step approach proposed by Almeida et al. (2022) for improving parametric implied volatility models. The researchers utilized neural network approximation to enhance the estimate of IVS derived from well-known parametric IV models. Neural networks possess the "universal approximators" property, allowing them to accurately approximate any Borel measurable function from one finite space to another. The authors successfully improved the IV estimates of Black-Scholes, Ad-hoc, Heston, Carr and Wu parametric models for options by reducing the pricing error surface with the neural network output. Furthermore, they investigated the Black-Scholes and Heston models when analyzing the options panel, a sequence of implied volatility surfaces, and enhanced the IV values using a non-parametric technique.

Although Almeida et al. (2022) only examined the Feed Forward Neural Network in their study, their research opens up opportunities for expanding this approach to include a broader range of neural networks. This paper aims to estimate IVS using well-known parametric models such as the Black-Scholes and Heston models, and subsequently, employ 1D convolutional neural networks (CNN), gated recurrent units (GRU), long short-term memories (LSTM), and feed-forward neural networks (FNN) for non-parametric correction. Therefore, the research question is: Can neural networks initially designed for time series outperform FNN in terms of non-parametric correction of IVS parametric estimates? By exploring the potential of neural networks in improving IV predictions, this study contributes to the ongoing research on the use of machine learning in financial markets and provides insights into the effectiveness of different neural network architectures for IVS prediction. Furthermore, the research focuses specifically on equity option IV, highlighting the importance of accurate IV predictions in stock options trading.

Our research delves into the comparison of different models for options pricing and their correction using neural networks. In our analysis, we observe that the Black Scholes model, which is a simple average implied volatility for a selected period, has a lower pricing error than the Heston model for American option pricing. This observation is significant for investors and traders as they rely on accurate options pricing to make informed investment decisions.

We further explore the effectiveness of non-parametric correction using neural networks. We find that neural networks that can capture time dependencies, such as LSTM, GRU, and CNN, show better non-parametric correction in general. However, LSTM exhibits more stable results than GRU and CNN, which may miss time patterns in the data and perform worse. CNN is generally better for Black Scholes correction, and GRU is better for the Heston model.

Incorporating time-varying covariates in the research can improve model correction. Surprisingly, for the Black Scholes and Heston models FNN, which are not trained to capture time-dependency, shows significant improvements of non-parametric correction. This finding suggests that incorporating relevant features can enhance the correction ability of the models. Among the time-dependent neural networks, we decided to utilize the LSTM network due to its stability in producing accurate results. Our findings indicate that the inclusion of features can lead to further improvements in correction accuracy for LSTM neural networks. Interestingly, we observe that even without incorporating features, LSTM still outperforms the FNN in terms of pricing accuracy. This suggests that the inherent capabilities of LSTM in capturing time dependency make it a strong candidate for option pricing modeling.

Furthermore, we investigate the impact of macro-factors on the pricing performance of the models. We find that including macro-factors such as the Aruoba-Diebold-Scotti business conditions index (ADS), Economic Policy Uncertainty (EPU), term spread, and credit spread worsens the pricing performance for all models in the stock options prediction exercise. However, for the selected American-type options, moneyness, left tail probability (LTP), left tail volatility (LTV) measures of Bollerslev et al. (2015), and the VIX index calculated for each stock individually can significantly improve the correction exercise.

Our research provides valuable insights for investors and traders who rely on options pricing for investment decisions. The use of two-step approach of Almeida et al. (2022) can significantly improve the accuracy of implied volatility estimations, thereby reducing pricing errors and improving profitability.

The paper is structured as follows: In Section 2, a review of relevant literature is presented. Section 3 provides an overview of the data, while Section 4 presents the methodology employed in this study. Finally, Section 5 reports the empirical results obtained from the analysis.

2 Literature Review

As implied volatility is a crucial parameter in option pricing and hedging, modeling IV has been a subject of research for decades. There are two major categories of methods for modeling IV: parametric and non-parametric. In parametric approaches, the implied volatility surface (IVS) is modeled by a small number of parameters that are fitted by market data and asset dynamics. Among parametric approaches, indirect models use a dynamic process to determine implied volatility, including Levy, local volatility, and stochastic volatility models. Some notable indirect parametric models include the Merton model (Merton (1976)), the Heston model (Heston (1993)), the Kou model (Kou (2002)), the Chockalingam model (Chockalingam and Muthuraman (2011)), and the Kenichiro model (Shiraya and Takahashi (2017)). Indirect models are simple to implement, but incorporating time-dependent characteristics increases processing complexity. In contrast to indirect models, direct methods set IV directly and can be categorized into two groups. The first group assumes that the IVS's dynamics change constantly over time (Cont and Da Fonseca (2002); Carr and Wu (2016)). The second group takes into account a static surface depiction of implied volatility and employs either parametric or non-parametric models for calculating IVS. The stochastic volatility (SVI) model presented by Gatheral (2004) is one such model. The research is modeled with a fixed repayment approach while considering volatility. And subsequently, Gatheral and Jacquier (2014) collectively expand this concept to provide a surface free of static arbitrage.

Recently, machine learning techniques have been extensively used to improve the accuracy of IV predictions. In the study of Horvath et al. (2021), deep neural networks were employed to optimize a significant number of stochastic volatility models, with IVs and prices represented as pixels. Cao et al. (2020) examined the connections between the anticipated daily change of the IV S&P 500 index and the index's daily return, the VIX fear index, time to maturity, and moneyness. The authors discovered that they could considerably enhance the performance of their analytical model by employing the VIX index and the daily returns of the index as the primary variables. For feature selection, Zhang et al. (2021) used three methods: principal component analysis (PCA), variational autoencoder, sampling the surface, and derived variable predictions using LSTM. In predicting IVS, some studies have used non-parametric approaches, such as the regression tree model in Audrino and Colangelo (2010), the long-term short-term memory model with an attention mechanism in Chen and Zhang (2019), and the temporal difference backpropagation (TDBP) model in Bloch and Böök (2020). Zeng and Klabjan (2019) created an IVS by employing a support vector regression model using high-frequency data alternatives. Meanwhile, Ning et al. (2021) suggested a method for modeling IVS without time-based arbitrage by using a variational autoencoder to produce the model parameters for the future date once the arbitrage-free stochastic model has been calibrated to the IVS data.

Most researches focus on modeling the implied volatility of European options. The valuation of American-type options is a captivating subject that has garnered significant attention in the financial industry due to its complexities. Unlike European options, American options can be exercised at any point in time prior to their expiration date. Consequently, pricing American options involves determining optimal exercise prices and option prices simultaneously, which renders the option pricing problem nonlinear (Yan et al. (2022)). As a result, obtaining analytical solutions for such problems has proven to be a challenging task. When it comes to determining the implied volatility of an American option, there are two popular approaches. The first method involves utilizing a de-Americanization technique, which essentially converts an American option price into its European counterpart (Burkovska et al. (2018) and Carr and Wu (2010)). However, this method has its limitations as significant pricing errors may occur due to the inaccurate incorporation of the early-exercise premium. In fact, the larger the early exercise premium, the higher the potential for errors to occur. Furthermore, incorporating dividends can further increase the margin of error of the de-Americanization method. The second approach involves a direct calibration process for the American-style pricing model to extract the implied volatility (Achdou et al. (2004), Lagnado et al. (1997)). In this approach, the implied volatility is computed by solving a minimization problem, using an iterative search technique that repeatedly compares a model with market option prices until a suitable value of volatility is found. Unlike European-style options, the derivative of the option value to the volatility does not have a closed-form expression for American-style options. Numerous numerical methods are used for American option pricing, including simulation-based techniques, binomial lattices, partial differential equation (PDE) solution methods, and integral methods. These techniques have been used extensively in prior research, with notable contributions by Longstaff and Schwartz (2001), Cox et al. (1979), Brennan and Schwartz (1977), Muthuraman (2008), Kim (1990), Jacka (1991), and Carr et al. (1992).

Despite the extensive research in this field, American options have not received as much attention as their European counterparts. Therefore, it becomes intriguing to explore non-parametric correction approaches, such as the one proposed by Almeida et al. (2022), to address the gap in implied volatility knowledge for equity options.

3 Data

In contrast to the approach taken by Almeida et al. (2022), this paper examines American call and put options. The present study investigates the predictive power of the implied volatility surface derived from the options of five leading tech companies - Amazon, Microsoft, Google, Facebook, and Apple (Figures: 1, 11, 12, 13). By analyzing their panel data, this research aims to shed light on IVS prediction. To obtain the IVS, we rely on Option Metrics data, which we rank over a period spanning from December 31, 2018, to December 31, 2021. Utilizing put-call parity is a commonly employed technique for inferring implied dividends. However, the put-call parity is not applicable to American-style options, as the presence of an early-exercise premium renders it invalid (Fodor et al. (2017)). Consequently, put-call parity only holds for European-style options and we do not utilize dividends in our research. As a benchmark for the risk-free rate, we download the 3-month Treasury bill rate from the St. Louis Federal Reserve Economic Data (FRED) database. To ensure the accuracy and reliability of our dataset, we adopt a rigorous methodology, in line with the established academic literature (Almeida et al. (2022); Andersen et al. (2015)).

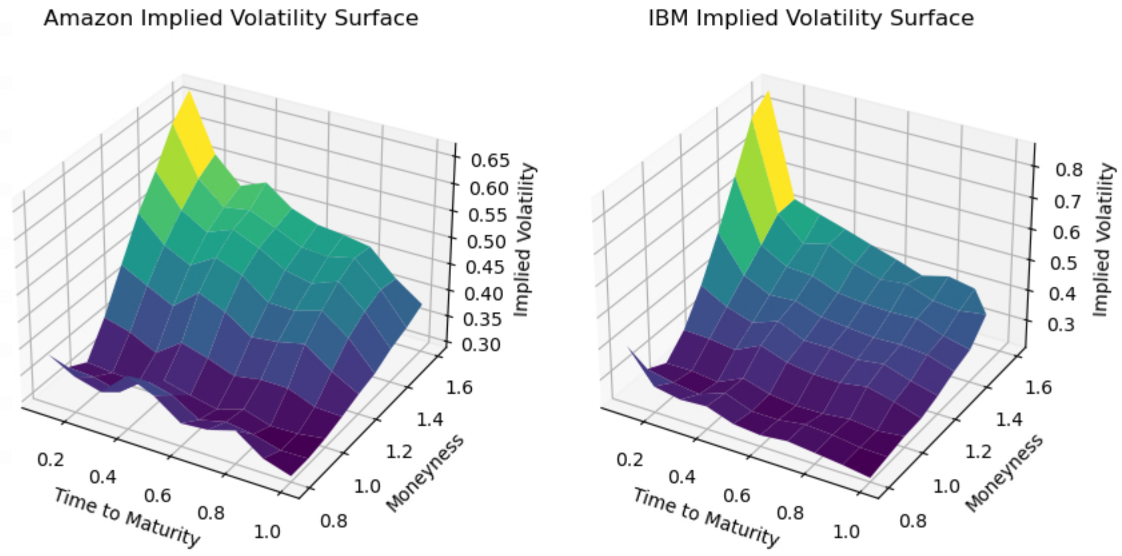

Figure 1: Implied volatility surface of Amazon and IBM.

Following the standard practice, we exclude in-the-money (ITM) options from our sample and focus exclusively on out-of-the-money (OTM) options, which are known to be more liquid and informative in the context of IVS analysis Andersen et al. (2015). Specifically, we divide call options into three categories, based on their moneyness: deep OTM, OTM, and close to ATM, with respective moneyness intervals of [0.8, 0.9), [0.9, 0.97), and [0.97, 1.03). Since the research focuses on OTM options, we excluded ATM options where the strike price equals the market price of the underlying option and left the category "close to ATM". For put options, we adopt a similar categorization, but with different moneyness intervals, namely OTM [1.03, 1.10) and deep OTM [1.10, 1.60). By selecting only short-term (20-60 days) and long-term (60-240 days) options, we ensure that our sample captures a broad range of market expectations and is not biased toward any specific time horizon. As we focus on American-style options, we do not impose any no-arbitrage constraints on our data. This choice allows us to investigate the full range of price dynamics available to market participants.

Looking at the data, we can see that among the selected stocks, Amazon stock options have the highest number of options totaling an impressive 463,985 after filtering. This figure is more than twice the number of options for any other dataset (Tables: 6, 7, 8, 9), indicating the popularity of Amazon stock among investors. However, while Amazon

Table 1: Summary of Amazon Stock Option Implied Volatility Statistics

Days to Maturity	Moneyness	Number	Mean IV	Std. dev. IV
Short	(0.8, 0.9]	25,310	0.345	0.120
Short	(0.9, 0.97]	40,519	0.308	0.102
Short	(0.97, 1.03]	25,106	0.299	0.090
Short	(1.03, 1.1]	39,477	0.330	0.100
Short	(1.1, 1.6]	60,002	0.438	0.148
Long	(0.8, 0.9]	42,449	0.349	0.120
Long	(0.9, 0.97]	54,977	0.314	0.085
Long	(0.97, 1.03]	32,856	0.303	0.083
Long	(1.03, 1.1]	50,927	0.323	0.084
Long	(1.1, 1.6]	91,909	0.412	0.143
Total	(0.8, 1.6]	463,985	0.355	0.126

This table provides a summary of the Amazon stock option implied volatility statistics for the period from December 31, 2018, to December 31, 2021. The table shows the count of options, the average implied volatility, and the standard deviation of implied volatilities for various time-to-maturity and moneyness categories. Short-term maturity options are considered as (20, 60] days to maturity, and long-term options as (60, 240], respectively.

options may be popular, they are also the most volatile in terms of implied volatility, with a value of 0.126. Higher implied volatility suggests that the market expects a greater degree of price movement in the underlying stock. Conversely, lower implied volatility indicates that the market expects less price movement in the underlying stock. On the other hand, the least volatile options in our dataset belong to IBM, with a standard deviation of the implied volatility equal to 0.105. This indicates that the market has lower expectations for price movement in IBM's stock, making it a potentially safer option for risk-averse investors. Interestingly, IBM options also have the lowest mean implied volatility compared to other stocks, with a value of 0.279. In contrast, other stocks have a mean implied volatility higher than 0.3. This indicates that IBM options are generally less volatile and may provide a more stable investment opportunity.

It is worth noting that while Amazon options are the most volatile, they do not have the highest mean implied volatility among all the stocks in our dataset. Apple stock options have the highest mean implied volatility at 0.364, indicating that the market expects significant price movements in Apple's stock over the coming period. When we look at the data more closely, we can see that the most volatile options in terms of implied volatility values are the category of short-term options, with a maturity of 20 to 60 days, and a moneyness range of 1.1 to 1.6. Moneyness is another important metric that measures the relationship between the strike price of an option and the current market price of the underlying asset. A call option with higher moneyness will have a higher market price of the underlying asset compared to the strike price, making it more likely to be in the money. In the case of at-the-money (ATM) call options, where the strike price is equal to the current market price, the moneyness is equal to 1. Moreover, we can observe that across all the stocks, the group of short-term implied volatility and moneyness from 0.97 to 1.03 has the least number of options. This suggests that investors may be less interested in options that are very close to being in the money or out of the money and are instead focusing on options that are further away from the current market price of the underlying asset. In conclusion, our data provide valuable insights into the options market for a selection of popular stocks. By analyzing implied volatility, and moneyness, time to maturity, we can gain a better understanding of market expectations for each stock.

4 Methodology

4.1 Error Correction Model

This research investigates a two-step approximation method proposed by Almeida et al. (2022) in the estimation of the implied volatility surface for equity options. The IVS is a fundamental tool in options pricing, mapping the implied volatilities σ to moneyness m and time to maturity τ. While several parametric models have been developed to improve the estimation of the IVS, they still face significant pricing errors. Therefore, Almeida et al. (2022)'s two-step approximation method, which combines both parametric and non-parametric techniques, and has proven potential to enhance the prediction performance of the IVS, is an interesting tool to apply for further research. In Almeida et al. (2022) paper, various parametric models were constructed, including the Black Scholes, Heston, ad-hoc Black Scholes, and Carr and Wu models of implied volatilities. Researchers extracted implied volatilities estimations of the parametric models and reduced the root mean squared error (RSME) between the observed implied volatilities of the S&P 500 index and the parametric IV estimations. In this study, we aim to extend the non-parametric approach and select the Black Scholes and Heston models as our parametric models. Given a n-sized cross-section of options $i = 1...n$ over a period of T days, the goal becomes to calibrate a parametric model p to the observed IVS $\sigma(m_{i,t}, \tau_{i,t})$, resulting in the fitted values

$\hat{\sigma}_p(m_{i,t}\tau_{i,t})$. The error function $\hat{\epsilon}_p(t, m_{i,t}, \tau_{i,t})$ is then computed by comparing the actual IVs with the estimates produced by the models (1), where i ranges over the moneyness and times to maturity for each day t.

$$\widehat{\epsilon}_p(t, m_{i,t}, \tau_{i,t}) = \sigma(t, m_{i,t}, \tau_{i,t}) - \widehat{\sigma}_p(t, m_{i,t}, \tau_{i,t}) \tag{1}$$

To improve the parametric performance of these estimates, we employ various neural networks, including FNN, LSTM, GRU, and 1D CNN, to minimize the error function ϵ_p for each neural network l as well as the pricing errors of the parametric models p. Our objective is to minimize the mean squared error function $\hat{\epsilon}_p$ (2) by selecting the neural network function $f(.)$ that minimizes the error most effectively.

$$\widehat{\epsilon}_l = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{n_t}\sum_{i=1}^{n_t}[\widehat{\epsilon}_p(t, m_{i,t}, \tau_{i,t}) - f(m_{i,t}, \tau_{i,t}))]^2 \tag{2}$$

Finally, the implied volatility surface is determined as the value of the IVS obtained from the parametric model, along with the adjustments made by non-parametric methods (3).

$$\widehat{\sigma}_l = \widehat{\sigma}_p(t, m_{i,t}, \tau_{i,t}) + \widehat{f}(m_{i,t}, \tau_{i,t}) \tag{3}$$

To evaluate the pricing performance of the combined models, we use the root mean squared error metric (Equation 4), which is calculated for both the parametric models and the BS-FNN, BS-CNN, BS-LSTM, BS-GRU, and Heston-FNN, Heston-LSTM, Heston-GRU, and Heston-CNN corrections. The models are then compared according to this metric to determine the best-performing one.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{4}$$

4.2 Parametric Models

In recent years, the financial modeling field has experienced a noteworthy transition towards utilizing machine learning algorithms. Among these algorithms, neural networks have emerged as a powerful tool for enhancing the precision of parametric implied volatility models, thus drawing considerable attention to their application. As a result, our research is primarily centered around exploring the potential of neural networks for refining two parametric implied volatility models: the Black-Scholes model and the Heston model. These models provide distinct approaches to comprehending implied volatility, and the objective of our study is to evaluate the efficacy of neural networks in improving their performance.

4.2.1 Black-Scholes Model

The Black and Scholes (1973) model is based on several key assumptions, including that the short-term interest rate (r) is known and constant over time. Additionally, it assumes that the stock price (S_t) follows a geometric Brownian motion with a constant drift (μ) and volatility (σ), and that the stock pays no dividends. Other assumptions include the absence of transaction costs, arbitrage opportunities, and the ability to borrow or lend any amount at the riskless rate (r). The model also assumes that it is possible to buy or sell any amount of the stock, including fractional shares.

$$\frac{\partial S_t}{S_t} = \mu dt + \sigma dW_t \tag{5}$$

In the formula 5, W_t represents a standard Wiener process, σ is constant instantaneous volatility, and μ represents expected asset return. Strong parametric assumptions allow the Black-Scholes model to produce a closed-form solution for the pricing of a European option. When utilizing the BS model, the following formula may be used to determine the price of a call option with the following parameters: price S, strike price K, and time to maturity τ.

$$\mathrm{C} = \mathrm{N}(\mathrm{d}_1)\mathrm{S} - \mathrm{N}(\mathrm{d}_2)\mathrm{K}\mathrm{e}^{-r\tau} \tag{6}$$

$$\mathrm{d}_1 = \frac{1}{\sigma\sqrt{\tau}}\left[\ln\left(\frac{S}{K}\right) + \tau\left(r + \frac{\sigma^2}{2}\right)\right] \tag{7}$$

$$\mathrm{d}_2 = \frac{1}{\sigma\sqrt{\tau}}\left[\ln\left(\frac{S}{K}\right) + \tau\left(r - \frac{\sigma^2}{2}\right)\right. \tag{8}$$

With the option's market price of the option, it is feasible to compute the value of IV using the Black Scholes model: $\sigma_i = C_{BS}^{-1}(C_i, S_t, K_i, \tau_i, r)$. Volatility being constant over time is one of the model's most fundamental assumptions. This is not compatible with reality, which is why the application of non-parametric approaches to the model will take the form of IVS into consideration.

The proposed approach can be viewed as an extension of the nonparametric estimation of the implied volatility surface. This can be demonstrated by noting that the following optimization problems are equivalent (Almeida et al. (2022)):

$$minf(\frac{1}{n}\sum_{i=1}^{n}[\sigma(m_{i,t},\tau_{i,t}) - f(m_{i,t},\tau_{i,t})]^2) \tag{9}$$

$$minf(\frac{1}{n}\sum_{i=1}^{n}[\sigma(m_{i,t},\tau_{i,t}) - c - f(m_{i,t},\tau_{i,t})]^2) \tag{10}$$

where c is any constant. Hence, a direct nonparametric estimation of the implied volatility surface can be interpreted as a correction of the Black-Scholes model, which predicts a flat surface when $c = \hat{a_0}$. However, since the Black-Scholes model does not provide any information about the shape of the implied volatility surface, correcting it is equivalent to fitting the surface directly.

4.2.2 Heston Model

The Heston (1993) model is one of the fundamental option pricing models used in finance. Like other stochastic models, the Heston model assumes that an asset's volatility follows a random process rather than a constant or deterministic one. In this model, the underlying asset price S_t is correlated with volatility, which follows the square root process V_t. The Heston model takes into account the return of volatility to the average value after reaching an extreme. The spot variance V_t is a key parameter in the model, which reverts to the long-run variance $\overline{v}$ at a rate m. The volatility of volatility, represented by σ_v, is also considered in the model. Additionally, the Heston model utilizes (11, 12) two Wiener processes, $W_{1,t}$ and $W_{2,t}$, to simulate the stochastic behavior of the underlying asset price and its volatility. The Heston model has been widely used in the financial industry due to its ability to capture the volatility smile and skew observed in options markets

$$\frac{dS_t}{S_t} = (r - \frac{1}{2}V_t)dt + \sqrt{V_t}dW_{1,t} \tag{11}$$

$$\mathrm{d}V_t = m(\overline{v} - V_t)dt + \sigma_v\sqrt{V_t}dW_{2,t} \tag{12}$$

To improve the accuracy of option pricing, various extensions of the Heston model have been proposed, including the inclusion of jumps, stochastic interest rates, and more complex forms of correlation. Despite these extensions, the Heston model remains a popular choice due to its simplicity and ability to accurately capture the main features of options markets. One of the main advantages of the Heston model is its ability to provide closed-form solutions for option prices. This is achieved through the use of Fourier transform techniques, which allows for fast and accurate computation of option prices. In addition, the model's parameters can be estimated using various techniques, such as maximum likelihood estimation and Bayesian methods. Despite its popularity and usefulness, the Heston model has some limitations. For example, it assumes that volatility is mean-reverting, which may not always hold in practice. In addition, the model can be computationally intensive, particularly when dealing with high-dimensional problems. Nevertheless, the Heston model remains a valuable tool for option pricing and risk management, and its extensions and modifications continue to be an active area of research in finance.

4.3 Non-parametric Models

ML algorithms, especially neural networks, are popular in financial modeling due to their effectiveness in improving the accuracy of implied volatility models. However, network architecture selection is critical for their performance. We chose several neural networks, including Feed Forward neural network, LSTM, GRU, and 1D CNN, for their ability to handle time series data effectively. LSTM and GRU are suitable for modeling sequential data and can capture long-term dependencies. 1D CNNs detect local patterns and can capture short-term fluctuations in the data. Utilizing these models, we make predictions on three-year option panel data, emphasizing the importance of temporal dependencies.

4.3.1 Feedforward Neural Network

The concept of FeedForward, which is a type of artificial neural network, was initially proposed by Rosenblatt (1958). FeedForward neural networks (FNN) have a layered architecture, with the first layer representing the input layer and the last layer representing the output layer. In their recent study, Gu et al. (2020) presented a framework that describes the architecture of FNN and estimated five models with various hidden layers. The first layer consists of 32 neurons, and the deep neural network's five hidden layers are composed of 32, 16, 8, 4, and 2 neurons, respectively. The structure of each hidden layer in a neural network can be described as a geometric pyramid rule. To stream information between the input and hidden layers, a set of units is used to determine the weighted sum of the neurons from the previous layer. The output layer is then activated through a nonlinear function. We adopt the exponential linear unit (eLU) functional form as the activation function for the inner layers as we are predicting errors that can take negative values. ELUs, unlike ReLUs, possess negative values that can drive the mean unit activations to be closer to zero, similar to batch normalization, but with reduced computational complexity (Clevert et al. (2015)). In the output layer linear activation function is utilized as it doesn't have boundaries and gives an opportunity to predict the most accurate value for the parametric error. To prevent overfitting in our FNN model, we employ regularization strategies such as learning rate shrinkage, as proposed by Kingma and Ba (2014), and the early stopping strategy, as described by Prechelt (1998). In addition to the regularization techniques, we

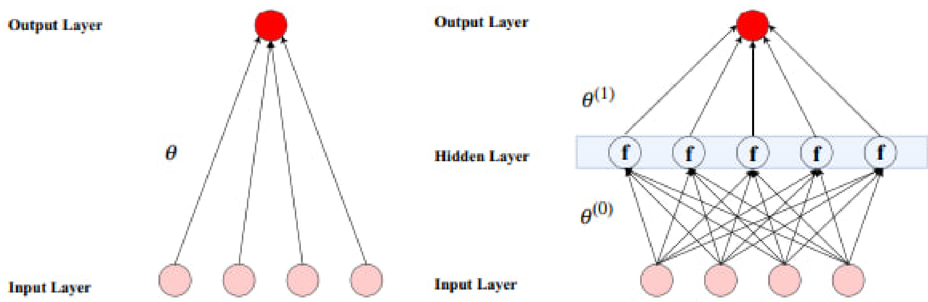

Figure 2: FNN architecture depicted in the figure (Gu et al. (2020)).

also investigate the adaptive moment estimation algorithm (Adam), a first-order gradient optimization method of stochastic objective functions, as developed by Kingma and Ba (2014). The Adam optimization technique is preferred over stochastic gradient descent (SGD) as it converges more quickly and requires less fine-tuning. As the FNN model needs to estimate a large number of parameters, the Adam optimization technique is an ideal choice. The learning rate, a tuning parameter for the Adam method, is determined for each parameter using the first and second-moment gradient estimations. We use the "learning rate shrinkage" proposed by Kingma and Ba (2014) to manage it.

4.3.2 Long-short-term-memory Neural Network

Long-Short Term Memory (LSTM) is a variant of Recurrent Neural Networks (RNN) that has gained widespread popularity in recent years due to its ability to learn long-term dependencies. RNNs are effective in handling sequential data but are limited by the vanishing and exploding gradient problem, which occurs when the gradients become too small or too large to be effectively propagated through time. LSTM solves this problem by introducing a gating mechanism that allows the network to selectively retain or forget information over time.

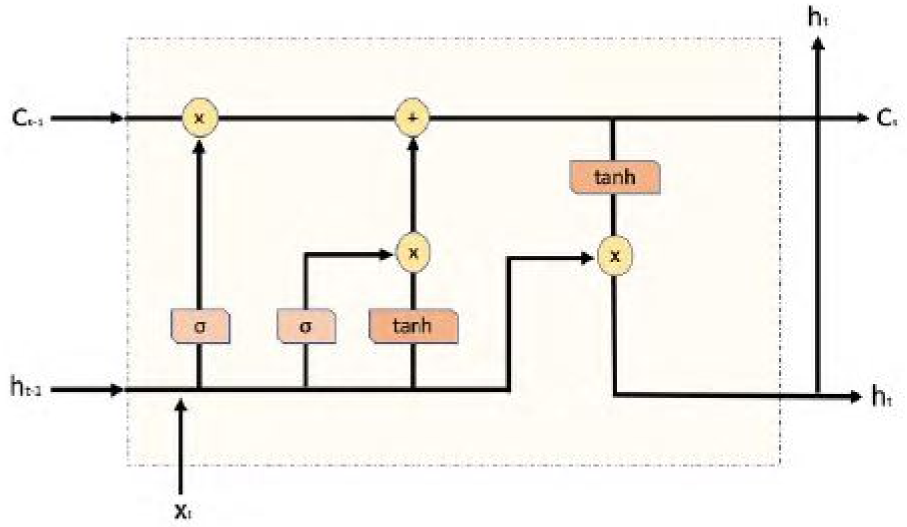

Figure 3: The LSTM architecture depicted in the figure (Ntakaris et al. (2019)).

The LSTM architecture was first introduced by Hochreiter and Schmidhuber (1997). Since then, many researchers have worked on improving and popularizing the technique. The core of the LSTM architecture is the memory cell, which stores information over time. Each neuron in the LSTM network has three gating mechanisms: the input gate, the output gate, and forget gate. These gates are controlled by sigmoid activation functions and act as switches that regulate the flow of information into and out of the memory cell. The input gate of an LSTM neuron determines which information from the current input and previous hidden state is relevant and should be stored in the memory cell. The forget gate, on the other hand, determines which information stored in the memory cell should be kept or discarded. Finally, the output gate controls the output of the cell and decides which information should be passed on to the next hidden state. Overall, the LSTM architecture has proven to be highly effective in a wide range of applications due to its ability to model complex, long-term dependencies in sequential data.

4.3.3 Gated Recurrent Unit Neural Network

As we already stated, the traditional RNN architecture suffers from the vanishing gradient problem, which hinders its ability to capture long-term dependencies. This issue prompted the development of more sophisticated RNN variants, such as the Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) architectures.

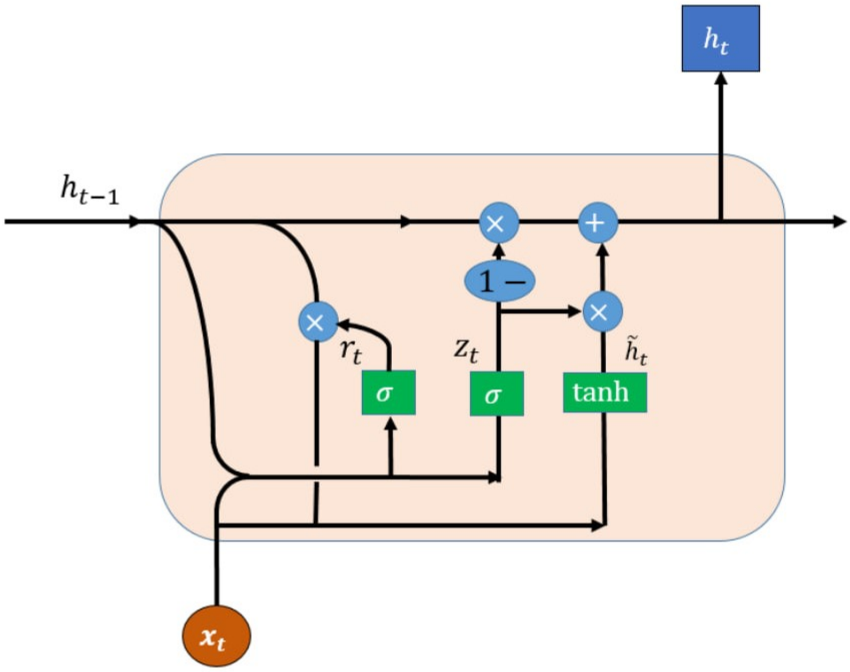

Figure 4: GRU cell architecture depicted in the figure (Huang et al. (2019)).

Introduced by Chung et al. (2014), the GRU is a promising alternative to the LSTM that offers a simpler architecture while still being capable of capturing long-term dependencies. Unlike the LSTM, which has three gates, the GRU has only two gates: the update gate and the reset gate. These gates act as switches that control the flow of information between the current input, previous hidden state, and current hidden state. The update gate determines how much information from the previous state should be passed on to the current state, while the reset gate determines how much information from the previous state should be ignored. By selectively passing information to the output based on the input and previous state, the GRU is able to effectively model sequential data while being computationally less expensive than the LSTM. By offering a more computationally efficient alternative to the LSTM, the GRU represents a valuable addition to the arsenal of deep learning practitioners.

4.3.4 1D Convolutional Neural Network

The 1D Convolutional Neural Network (CNN) is an architecture that has revolutionized pattern recognition tasks and is not often using particularly with time-series data, opening opportunities for further research. It has a unique structure consisting of three main layers, namely the convolutional layer, the pooling layer, and the fully connected layer. This neural network was first proposed by LeCun et al. (1989), and since then, it has evolved to become a powerful tool for researchers and practitioners alike.

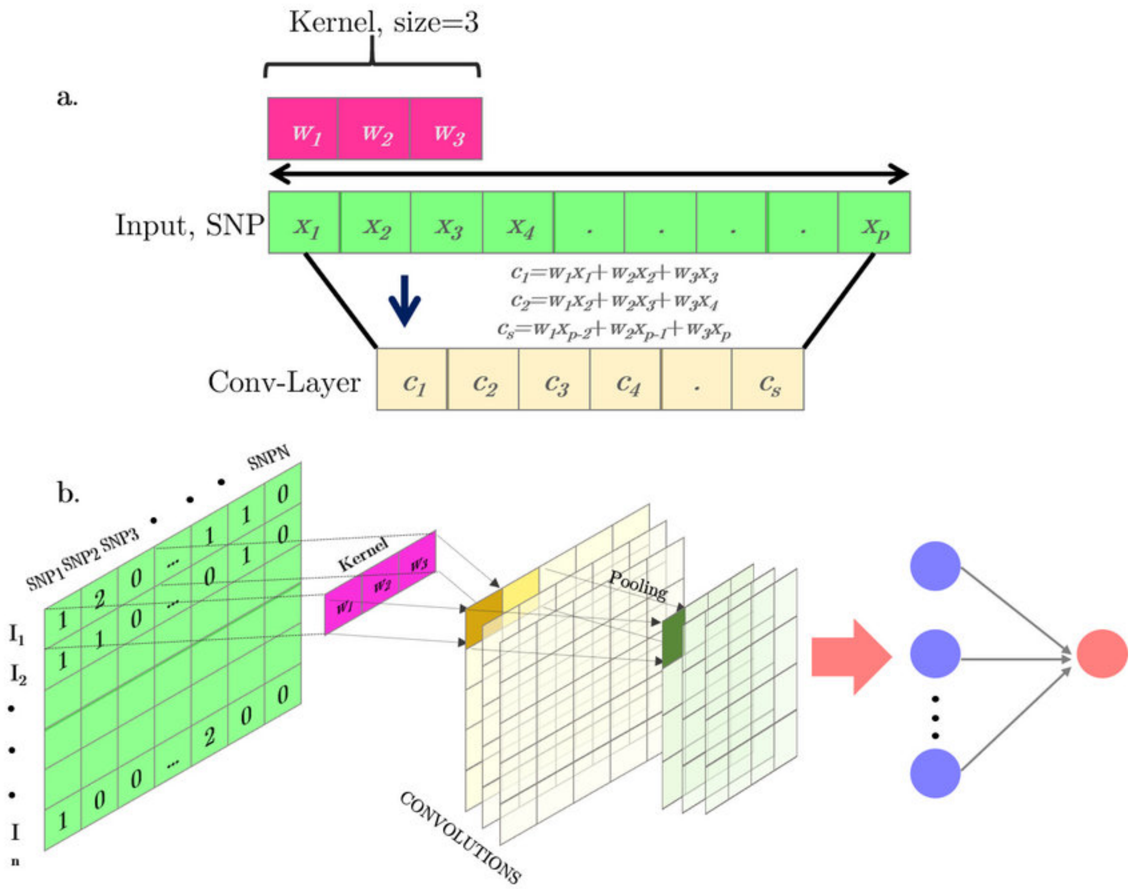

Figure 5: The CNN 1D architecture depicted in the figure (Pérez-Enciso and Zingaretti (2019)).

The convolutional layer applies a mathematical operation known as convolution to the input data using a set of filters or kernels. These kernels are essentially small matrices of weights that slide over the input data and produce a feature map that highlights specific patterns or features in the input data. The pooling layer then downsamples the feature maps by reducing their spatial resolution, increasing the network's translational invariance, and retaining only the most critical features. Finally, the fully connected layer applies a non-linear activation function to the output of the pooling layer and performs classification or regression on the resulting feature vector.

In 1D CNNs, the convolution operation is applied only in one direction, typically along the time axis in time-series data. This unique property of 1D CNNs allows them to capture temporal dependencies in the data, making them particularly useful for analyzing time-series data. By learning features that are invariant to changes in time, 1D CNNs can effectively identify trends and patterns in time-series data that would be challenging to detect using other traditional methods.

4.4 Implementation Details

4.4.1 Parametric Models

In this study, we utilize implied volatility data from Option Metrics to estimate the Black Scholes implied volatility and compare it to the non-parametric approach. Our findings show that the Black Scholes implied volatility is just the average of the historical data, consistent with the non-parametric approach (4.2.1). To estimate the Heston model parameters, we analyze a sequence of cross-sectional data over a period of T days, with n_t observations each day. These observations include implied volatilities, strikes, and times to maturity. We use the non-parametric spot variance estimator proposed by Todorov (2019) and estimate the structural parameters of the Heston model on the training dataset denoted by $\xi = (v, \kappa, \sigma_v, \rho)$ by minimizing the sum of squared differences between the observed implied volatilities and those predicted by the Heston model for each day and each observation. This can be expressed mathematically as $\sum_{t=1}^{T} \sum_{j=1}^{n_t} [\sigma_{j,t} - \sigma_H(\xi, \hat{V}_t, S_t, K_{j,t}, \tau_{j,t}, r_t)]^2$. Once we find the constant structural parameters, we estimate the implied volatility from Heston model prices.

4.4.2 Non-parametric Models

The dataset used for our study consists of 2 years of implied volatility surface (IVS) data for each stock option in the training set and 1 year of data for testing. This proportion of data is utilized for all the parametric and non-parametric models in our study. To further lower the MSE of our predictions, we incorporate a validation set, which is equal to 20% of the training dataset, and use the MSE and early stopping techniques on the validation MSE. The use of a validation set helps in reducing overfitting, and the early stopping technique helps in stopping the training of the model when the validation error starts to increase.

In order to tune the hyperparameters of our neural network, we use a wide range of values for each hyperparameter. Following the approach used in Gu et al. (2020), we employ a geometric rule for the hidden layers (4.3.1), and tune the number of hidden layers from 1 to 5. Additionally, we tune the learning rate from 1e-1 to 1e-3, and set the batch size in the neural network to a constant value of 128. To reduce the mean squared error (MSE) of our predictions, we also tune the dropout rate from 0.1 to 0.5, and select the epoch value from a list of 200, 300, or 500 epochs. To increase the predictive power of our model, we introduce L1 regularization and tune the penalty from 0.001 to 1. In total, there are 21 hyperparameters that are tuned via the Optuna library (Akiba et al. (2019)). By automating the process of hyperparameter tuning, Optuna can help improve the performance and generalization of machine learning models, reduce the time and cost of model development, and facilitate the reproducibility and transparency of research results. The number of trials in the Optuna library is set to 100. The optimized hyperparameters allow us to obtain a more accurate and robust neural network model for predicting implied volatilities.

For the LSTM and GRU models, we consider a range of hyperparameters to optimize their performance. Specifically, we use 8 time periods in the LSTM and GRU architecture for training, we tune the number of LSTM/GRU units from 32 to 64, the number of dense units between 32 and 64, the dropout rate from 0 to 0.5, and the learning rate from 0 to 0.5. We utilize the Optuna library to efficiently tune the hyperparameters of our models, and due to computational constraints, the number of trials is set to 10.

For the CNN model, we set the number of epochs to 100, and select 8-time steps for prediction. We then tune the number of convolutional layers between 32 and 64, kernel sizes between 3 and 5, and the number of dense units between 32 and 64. Additionally, we tune the dropout rate from 0 to 0.5, and the learning rate from 0.0001 to 0.1. To efficiently search the hyperparameter space, we utilize the Optuna library with 50 trials.

The selection of optimal hyperparameters for our models plays a crucial role in the performance of our neural network models. The use of the Optuna library allows us to efficiently search the hyperparameter space (Table: 10) and obtain an optimal set of hyperparameters for our models. By tuning the hyperparameters of our FNN, LSTM, GRU, and CNN models, we are able to improve the predictive accuracy and robustness of our models and obtain better results for predicting implied volatilities.

4.5 Time-Varying Covariates

The present study aims to enhance the predictive accuracy of our model by incorporating time-varying features. As financial markets are highly dynamic and ever-changing, it is imperative to consider the changes in covariates over time for making accurate forecasts. Following the approach of Almeida et al. (2022), we have included seven time-varying variables in our analysis.

One of the most significant time-varying variables that we have included in our study is the CBOE Volatility Index (VIX), also referred to as the "fear index." The VIX serves as a crucial leading indicator of the U.S. stock market, reflecting market participants' expectations of volatility over the next 30 days. It is a vital benchmark of market sentiment and risk, and its inclusion in our analysis could enhance the accuracy of our forecasting model. We have incorporated the CBOE VIX index for all stocks except Microsoft, for which we have used the standard VIX index. Since the VIX index captures both diffusive and jump risks, we have also included a separate measure of market jump risk developed by Bollerslev et al. (2015). The method involves calculating the likelihood that the index will fall by 10% or more over the following week (left tail probability or ltp) as well as the predicted (risk-neutral) return volatility caused by substantial negative price surges (left tail volatility or ltv). We have relied on Bollerslev et al. (2015) research to calculate both left-tail volatility and left-tail probability for each stock in Matlab. Additionally, we have incorporated realized volatility, which measures the historical variation in a security's price over a specified period, using 10-day realized volatility data from OptionMetrics.

Apart from time-varying variables, we have also incorporated macroeconomic factors into our model. Similar to Almeida et al. (2022) approach, we have used the Economic Policy Uncertainty (EPU) index developed by Baker et al. (2016), which draws on newspaper archives from Access World NewsBank service. The EPU index captures the level of uncertainty about economic policy that may affect future economic activity. By incorporating the EPU index into our model, we could capture the impact of economic policy changes on stock prices. We have also included the Aruoba-Diebold-Scotti (ADS) business conditions index, published by the Federal Reserve Bank of Philadelphia, in our analysis (Aruoba et al. (2009)). The ADS index provides a summary of the current state of the U.S. economy, offering insight into the direction and amplitude of the business cycle. The inclusion of the ADS index in our model could help us capture the impact of macroeconomic factors on stock prices. Furthermore, we have considered the first differences between the term spread and credit spread features. The term spread reflects the difference between the 10-year Treasury rate and the 3-month Treasury rate, while the credit spread indicates the difference between the yield on Moody's Seasoned Baa Corporate Bonds and the yield on the 10-year Treasury. The incorporation of these features could help us capture the impact of interest rates on stock prices.

To assess the impact of features on the output, we employ an altered version of the permutation feature importance technique. The concept of feature importance has gained immense popularity in recent years due to its ability to provide insights into the relative contributions of different factors toward the prediction of a target variable. In particular, we use the feature importance metric of Gu et al. (2020). The I_{in} is determined by the change in the RMSE metric resulting from lowering all values of feature n to zero while keeping the other model estimates and features constant. The RMSE feature importance is a powerful tool for identifying significant features that aid in the prediction of implied volatilities. A feature that has a low I_{in} can be considered insignificant since its removal does not significantly affect the predictive performance of the model. Conversely, a feature with a high I_{in} is a strong indicator of its importance in improving the accuracy of the model. This information is critical for model interpretation, variable selection, and ultimately, the development of more accurate predictive models.

Our research focuses on utilizing well-established characteristics of the implied volatility surface to explore the predictive ability of implied volatility in capturing its non-linear dynamics (Andersen et al. (2015). These characteristics have been studied in Almeida et al. (2022) specifically and include the level, term structure, skew, and skew term structure. The level of the implied volatility surface pertains to the average implied volatility of short-term at-the-money options. On the other hand, the term structure represents the difference between the average implied volatility of long- and short-term ATM options. The skew is defined as the difference between the average implied volatility of short-term OTM put and call options. Finally, the skew term structure is the difference between the long- and short-term skew, where the long-term skew is defined in a similar manner as the short-term skew. We evaluate the root mean square error (RMSE) of the level, term structure, skew, and skew term structure by comparing the observed implied volatility characteristics with the predicted ones. This approach enables us to assess the accuracy of our predictions in summarizing the behavior of the implied volatility surface.

5 Empirical Results

5.1 Option Panel Performance

The study is structured into two distinct segments: namely the assessment of performance in the option panel and the examination of performance in the presence of time-varying covariates. In this section, we elaborate on the findings of the former, particularly on the options panel performance, which involves a thorough analysis of the time to maturity and moneyness effect on parametric error. Additionally, we delve into the performance of options that have been categorized based on moneyness and time to maturity, thereby providing a more nuanced understanding of the efficacy of the models employed in this study.

The research focused on estimating implied volatility for American options using various models. The results (Table: 2) showed that a historical average can be a better estimation of implied volatility for American options. The Black-Scholes model outperformed the stochastic Heston model for every stock. Most research on implied volatility forecasting of American-type options is performed mostly under constant volatility and there are fewer studies under stochastic volatility due to the nature of equity options (Chockalingam and Muthuraman (2011)). The superior performance of the constant volatility model over a stochastic one for stock options was identified in the current research. However, the non-parametric correction showed its effectiveness for both Black Scholes and Heston implied volatility models. Furthermore, introducing neural networks improved the performance of both parametric models.

Table 2: Summary of Black Scholes and Heston model corrected performance

	Black Scholes					**Heston**				
	No NN	**FNN**	**LSTM**	**GRU**	**CNN**	**No NN**	**FNN**	**LSTM**	**GRU**	**CNN**
AAPL	0.089	**0.066**	0.067	0.068	**0.066**	0.119	0.097	0.069	0.064	**0.061**
AMZN	0.112	0.100	0.069	0.073	**0.066**	1.670	0.583	**0.464**	1.155	1.185
GOOGL	0.100	0.084	0.060	0.060	**0.056**	0.182	0.107	**0.081**	0.121	0.122
MSFT	0.087	**0.062**	**0.062**	0.079	0.063	0.149	0.063	0.065	**0.057**	0.064
IBM	0.073	0.067	**0.059**	0.127	0.122	0.199	0.091	0.071	**0.065**	0.122

The table provides a summary of the pricing performance of the FNN, LSTM, GRU, and 1D CNN neural networks combined with Black Scholes and the Heston model measured in RMSE (root mean squared error). The pricing performance is measured from 31 December 2020 to 31 December 2021.

Concluding from the combined results, the parametric models were outperformed by neural network correction. CNN neural network showed the lowest RMSE of implied volatility estimation for Google options for the Black Scholes model. Regarding the Heston model, the best-performing model in terms of RMSE turned out to be GRU neural network, applied to Microsoft options prices. However, across both Black Scholes and

Heston correction time-series-dependent neural networks did not always outperform FNN, which is less computationally expensive and nevertheless provided decent results.

It is worth noting that The 1D CNN Black Scholes model consistently showed lower estimation error than the FNN benchmark, but performance seems to deteriorate for small datasets such as IBM options, where GRU neural networks struggle too. We can see that GRU neural network for the Heston model showed very consistent results, outperforming benchmarks, FNN, and LSTM in four out of five instruments, only falling behind with Google options. Overall, across both models and different instruments, GRU and CNN models are performing very well, drastically reducing the RMSE of the estimation in most cases. We can see that LSTM performed as a very stable neural network, providing RMSE reduction every single time, however, often falls behind GRU and CNN in terms of the magnitude of the provided reduction.

Table 3: Black Scholes and Heston model corrected performance for Amazon stock options

	Black Scholes					**Heston**				
	No NN	**FNN**	**LSTM**	**GRU**	**CNN**	**No NN**	**FNN**	**LSTM**	**GRU**	**CNN**
All	0.112	0.100	0.069	0.073	**0.066**	1.670	0.583	**0.464**	1.155	1.185
DOTMC	0.114	0.114	0.079	**0.070**	0.078	1.721	**0.593**	0.830	1.250	1.201
OTMC	0.093	0.077	**0.061**	**0.061**	**0.061**	1.940	0.178	**0.211**	1.674	1.673
ATM	0.062	0.064	0.054	**0.053**	0.054	2.524	0.214	**0.160**	2.458	2.432
OTMP	0.070	0.075	0.061	**0.060**	0.061	0.332	0.334	**0.234**	0.253	0.227
DOTMP	0.135	0.123	0.090	0.083	**0.080**	0.419	0.342	**0.220**	0.304	0.261
Short	0.108	0.089	**0.068**	0.069	**0.065**	2.021	**0.495**	0.501	1.607	1.513
Long	0.115	0.108	0.070	**0.068**	0.069	0.326	0.301	**0.206**	0.299	0.234

The table provides a summary of the pricing performance of the FNN, LSTM, GRU, and 1D CNN neural networks combined with Black Scholes and the Heston model measured in RMSE (root mean squared error) per category. The definition of each category can be found in 3.The pricing performance is measured from 31 December 2020 to 31 December 2021.

If we study the non-parametric correction by categories in Table:3, we can notice the following results. For all options data categories, the CNN and LSTM models performed best in estimating option prices for the Black Scholes and Heston models, respectively. Notably, without the inclusion of neural networks, the difference in RMSE between the Black Scholes and Heston models is striking, with the latter exhibiting RMSE values that are four to ten times higher. This is in contrast to previous research conducted by Almeida et al. (2022) which found that the Heston model outperformed the Black Scholes model for

European-type options. Our results indicate that the Black Scholes estimation of implied volatility (IV) through averaging across periods is a better estimation of IV for stock options than a stochastic Heston model for all categories. Furthermore, introducing a non-parametric correction can also significantly reduce RMSE for each category. However, even with this correction, the Black Scholes model still outperforms the Heston model. Interestingly, the neural network correction did not improve the parametric models to the same extent.

The GRU neural network achieved the absolute minimum RMSE of 0.054 when applied to the Black Scholes model in the at-the-money option category. In the on-the-money call options and deep on-the-money call options categories, LSTM, GRU, and CNN neural networks showed the greatest reduction in RMSE for the Black Scholes model. These models reduced the error rate to the range of 0.06 to 0.08. On the other hand, FNN and LSTM neural networks exhibited the most consistent results for the Heston model on-the-money call options. The FNN neural network achieved the minimum RMSE of 0.593 for deep out-of-the-money calls in the Heston model, which represents a reduction in RMSE of almost 66%. Our analysis also considered options with short and long maturities. LSTM, GRU, and CNN models were found to perform well when applied to the Black Scholes model, achieving a significant reduction in RMSE. Conversely, the best performers for the Heston model were the FNN and LSTM models. In the at-the-money options category, the GRU neural network achieved the absolute minimum RMSE for the Black Scholes model.

However, both GRU and CNN models failed to achieve a significant RMSE reduction for the Heston model. Across on-the-money put and deep put options, LSTM, GRU, and CNN performed fairly well, showing consistent RMSE reductions in both models. In terms of comparing the models, we found that the GRU and CNN models dominated in terms of RMSE reduction for the Black Scholes model. In contrast, the LSTM model exhibited the best performance for the Heston model, dominating five out of eight categories. The FNN model also performed well, achieving the best scores in the remaining two categories. The tables for other stock options can be found (11, 12, 13, 14). We receive similar results, where LSTM showed stable performance, whereas GRU and CNN models showed unstable results, showing the best results across categories and worse results compared to the FNN for the small datasets. FNN was the leading performer in some option categories. Overall, our results highlight the importance of considering different neural network models when estimating option prices and underscore the superior performance of the Black Scholes model correction compared to the Heston model for equity options.

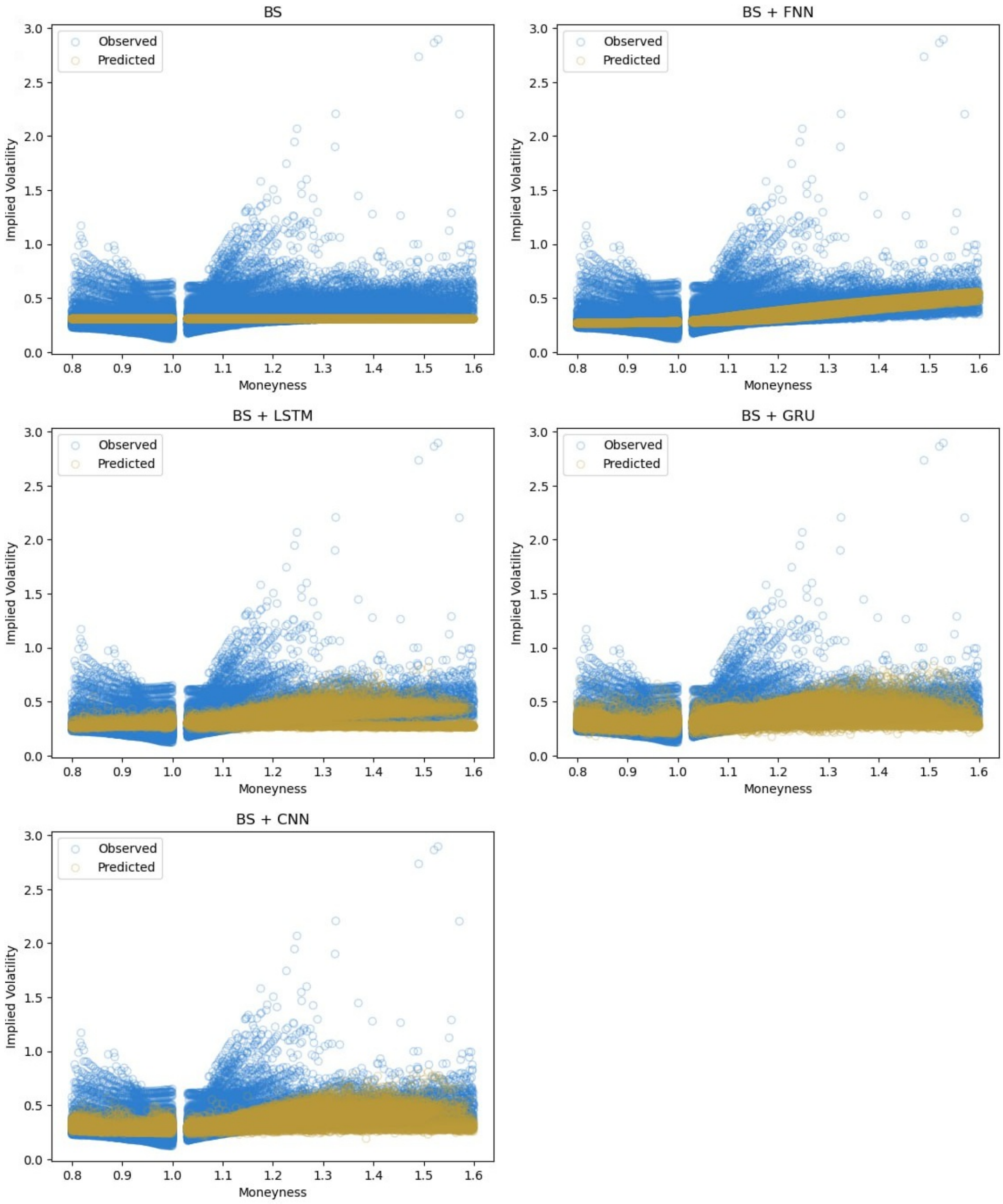

Figure 6: Prediction of Black Scholes, Black Scholes-FNN, Black Scholes-LSTM and Black Scholes-GRU in the Google option panel from 31 December 2020 to 31 December 2021.

In figures 14, 15, 16, 17, 18, 19, 20, 21 we examine the observed and predicted implied volatility values across various stock options, using data from the test period between December 31, 2020 and December 31, 2021. The Black Scholes model predicts constant implied volatility across all options, which is consistent with its theoretical assumptions. However, Black Scholes corrected results capture the observed pattern of the implied volatility smile slightly flatter than Heston's correction. This is in contrast to the findings of Almeida et al. (2022), who reported that the non-parametrically corrected Heston model outperformed the FNN Black Scholes model in generating the implied volatility smile, which is characterized by a U-shaped curve.

Our analysis reveals that the GRU and CNN corrected models exhibit a bias in their predicted IV values compared to the observed IV. In particular, the implied volatility smiles for Google stock options (Figure: 16) show a clear difference between the corrected models and the observed IV. We observe that the time-dependent neural networks used for the Black Scholes correction generate a wider range of predicted IV values compared to the FNN, while still partially resembling the implied volatility smile. For the Heston model correction, the pure Heston model and its corrections also do not fully reproduce the implied volatility smile (17). This highlights the limitations of the Heston model in capturing the complexity of American-type options IV and the need for alternative stochastic approaches for it.

5.2 Time-Varying Covariates Performance

In our study, we sought to investigate the impact of time-varying features on the accuracy of IV predictions. To that end, we employed neural network correction models and observed that incorporating the selected list of time-varying features led to worse results compared to models without such features. However, upon deeper investigation, we found that when macroeconomic features were included, all features had zero importance values. This led us to conduct an iterative process of exclusion, removing macroeconomic features one by one, until non-zero feature importance values were revealed for all the stocks options IV predictions when we omitted EPU, ADS, term spread, and credit spread.

Our results suggest that macroeconomic data does not necessarily improve the prediction of IV for stock-based options, which is consistent with previous research Almeida et al. (2022). To further explore this topic, we displayed the feature importance of Apple stock in 11. Interestingly, our findings revealed that moneyness is the most crucial feature for both Black Scholes and Heston model correction. The left tail probability and left tail volatility followed as the second and third most important features for Heston correction, while ltv and VIX index were the second and third most important features for the Black Scholes model. We also observed that feature importance values were higher for Black Scholes correction than for Heston, indicating that features have a more significant impact on Black Scholes prediction than in the stochastic Heston model. Furthermore, for stock

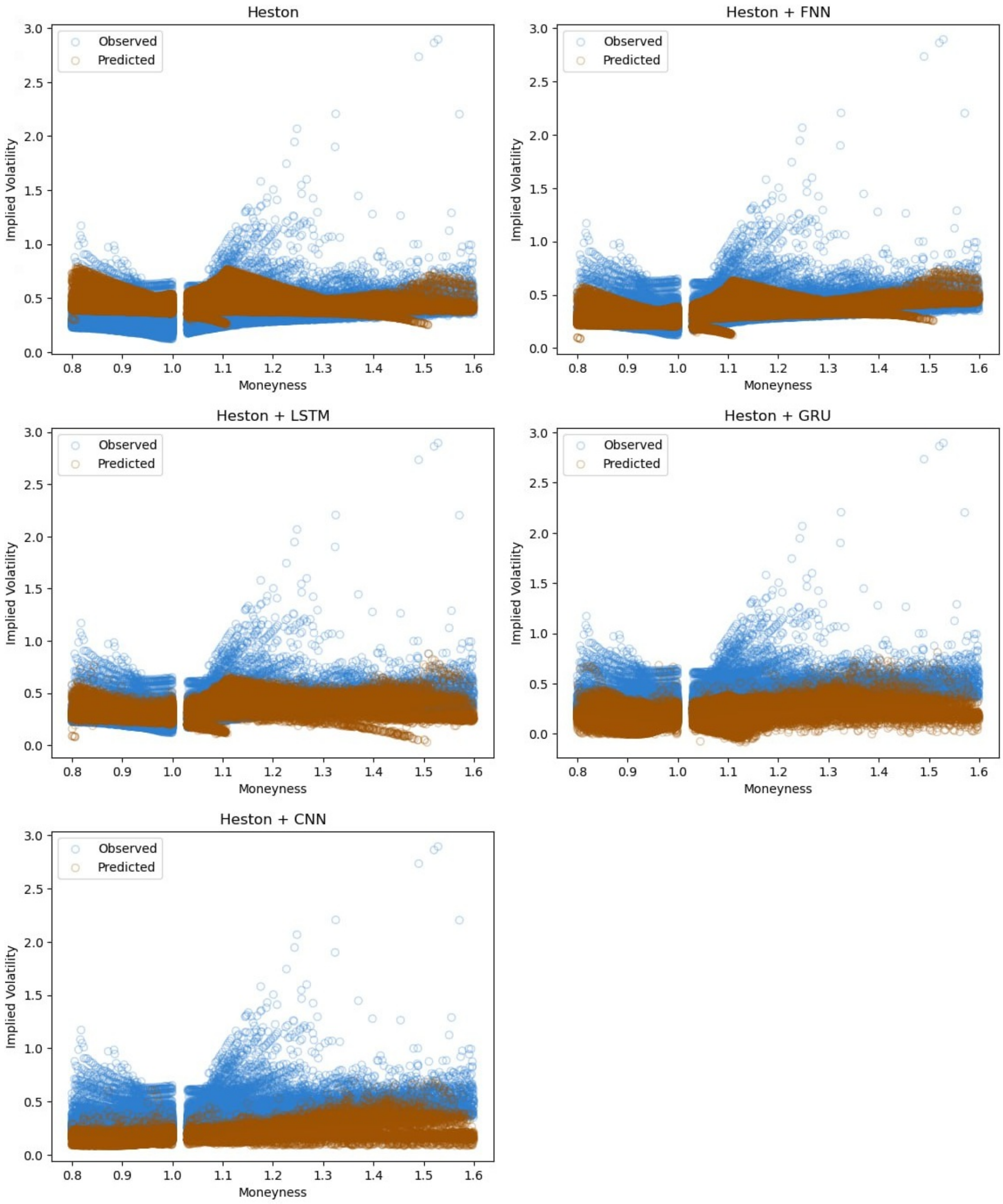

Figure 7: Prediction of Heston, Heston-FNN, Heston-LSTM and Heston-GRU in the Google option panel from 31 December 2020 to 31 December 2021.

options, features affect the more challenging prediction model (Heston) compared to the better-predicted average estimation of IV. Overall, our findings highlight the importance of carefully selecting relevant features for stock-based option IV prediction and suggest that incorporating macroeconomic data may not necessarily improve the accuracy of prediction models.

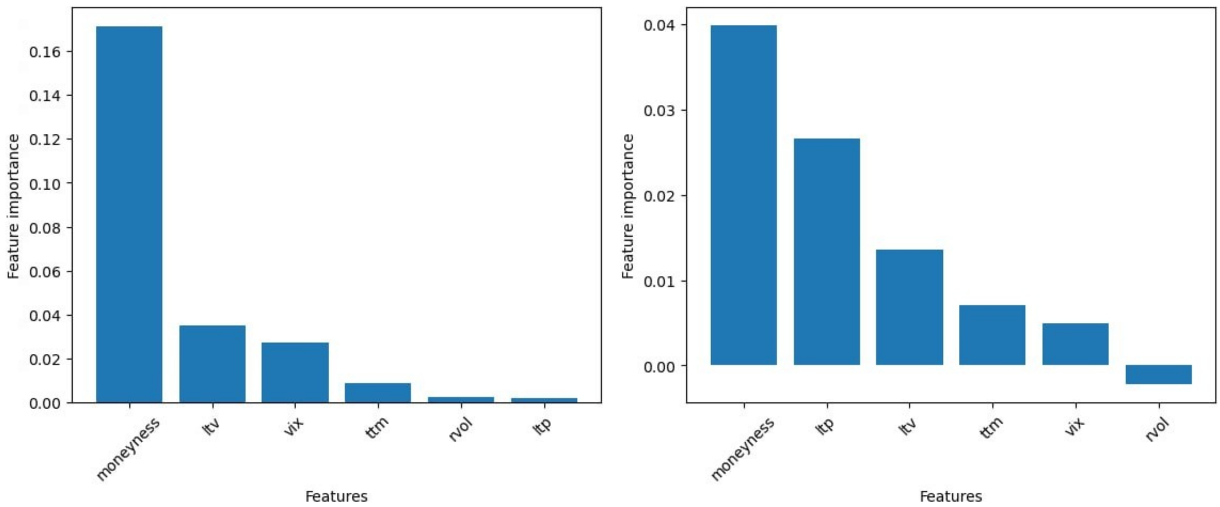

Figure 8: Feature importance of Apple Stock correction of Black Scholes and Heston model from 31 December 2020 to 31 December 2021. Features include: moneyness, ltv, ltp, vix, time to maturity (ttm), and realized volatility (rvol), which definitions can be found in Section 3.

To assess the performance of various models for time-varying covariates, we selected FNN, FNN with features (FNN+F), LSTM, and LSTM with features (LSTM+F). The FNN serves as the benchmark model, and LSTM is the most robust time-dependent neural network. The selection of LSTM is based on our desire to achieve superior prediction results on average. Our analysis of the Apple stock performance reveals that FNN based model with features outperforms other models for Black Scholes correction, exploiting all data variables present in the Apple stock. Meanwhile, LSTM without features outperforms models with features for all options for the Heston model.

However, when we analyze the options categories individually, we notice that FNN with features, LSTM, and LSTM with features models perform better than others for Black Scholes correction. FNN without features does not perform better than the other models. Moreover, we can conclude that models that incorporate time series characteristics through their structure (as LSTM) or time-varying covariates (as FNN with features) perform better in the option panel. For the Heston model, we observe the superior performance of FNN with features among the categories, leading to an intriguing insight that a model which does not take into account time-series patterns by its construction as FNN with features outperforms the LSTM neural network, which is explicitly built for time series data. Overall, our findings indicate that the inclusion of time-varying covariates significantly improves FNN's performance, while it only provides slight improvements in time-dependent neural

Table 4: Black Scholes and Heston model corrected performance of time-varying features for Apple stock options

	Black Scholes					Heston				
	No NN	**FNN**	**+ F**	**LSTM**	**+ F**	**No NN**	**FNN**	**+ F**	**LSTM**	**+ F**
All	0.088	0.066	**0.055**	0.067	0.071	0.119	0.097	0.071	**0.069**	0.081
DOTMC	0.066	0.066	0.043	**0.052**	0.062	0.121	**0.034**	0.048	0.251	0.115
OTMC	0.067	0.068	0.061	**0.055**	0.073	1.939	1.938	**0.058**	1.739	0.096
ATM	0.073	0.074	**0.062**	0.063	0.335	2.441	0.157	**0.066**	0.412	0.099
OTMP	0.068	0.138	0.065	0.064	**0.053**	0.334	0.390	0.078	0.252	**0.072**
DOTMP	0.088	0.072	0.071	0.071	**0.068**	0.379	0.285	**0.075**	0.255	0.077
Short	0.101	0.091	**0.066**	0.075	0.075	1.919	1.085	0.098	1.252	**0.090**
Long	0.082	0.078	**0.055**	0.065	0.067	0.232	0.164	0.073	0.131	**0.073**

The table provides a summary of the pricing performance of the FNN, FNN + F, LSTM, and LSTM + F combined with Black Scholes and the Heston model measured in RMSE (root mean squared error) per category. The pricing performance is measured from 31 December 2020 to 31 December 2021.

networks. Therefore, we can conclude that our research demonstrates the effectiveness of incorporating time-varying covariates in implied volatility forecasting.

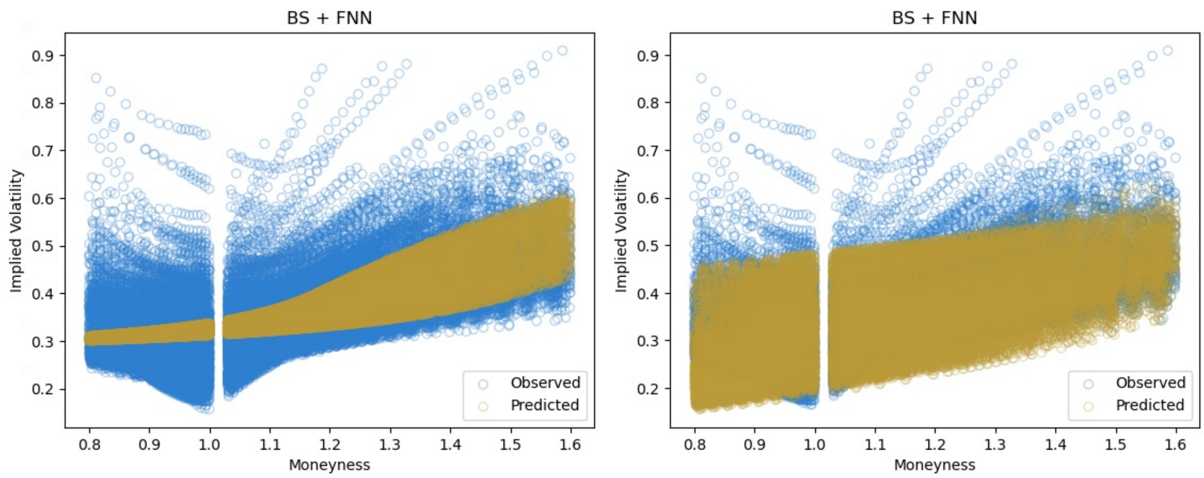

Figure 9: Prediction of Black Scholes FNN and Black Scholes FNN after including time-varying covariates from 31 December 2020 to 31 December 2021.

In Figures 9 and 10 we can see that after introducing the features into the neural networks, the range of implied volatility values becomes wider for both Black Scholes and Heston model correction exercise.

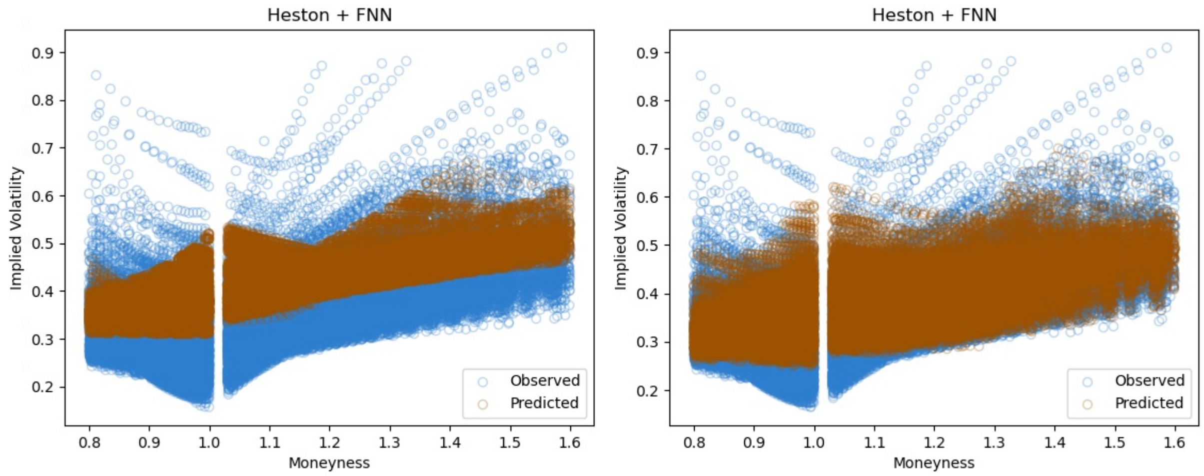

Figure 10: Prediction of Black Scholes FNN and Black Scholes FNN after including time-varying covariates from 31 December 2020 to 31 December 2021.

Table 5: Implied volatility characteristics of Apple option for Black Scholes and Heston corrected models.

	BS	**+FNN**	**+ F**	**+LSTM**	**+ F**	**H**	**+FNN**	**+ F**	**+LSTM**
Level	0.053	0.019	0.014	0.042	**0.013**	0.145	0.067	0.062	**0.034**
TS	0.003	**0.001**	0.004	0.011	0.002	0.041	0.041	0.051	0.004
Skew	0.091	**0.016**	0.095	0.152	0.095	**0.090**	0.095	0.094	0.092
Skew TS	0.012	**0.003**	0.015	0.029	0.011	0.007	**0.002**	0.012	0.011

The table provides a summary of the RSME of implied volatility characteristics of FNN, FNN with features, LSTM, and LSTM with features combined with Black Scholes and the Heston model. The measures are calculated from 31 December 2020 to 31 December 2021.

When we analyze the RSME of implied volatility characteristics, we reveal that the LSTM neural network, which was chosen as the benchmark for time-dependent neural networks, is outperformed by the FNN correction method that incorporates time-varying covariates. The surprising outcome of this study is that the neural network trained to learn from time patterns failed to outperform the Black Scholes and Heston FNN models with features in terms of level, term structure, skew and skew term structure characteristics. This observation is supported by the data presented in Table 5.

6 Conclusion

In this research, we extend the nonparametric correction of implied volatility models introduced by (Almeida et al. (2022)). Our goal is to investigate the performance of American-type options for major technology companies, including Amazon, Apple, Google, Microsoft, and IBM. Unlike Almeida et al. (2022), who focused on European options, we apply our methodology to American options. After filtering our data, we estimate both the Black-Scholes model and the Heston model. Our analysis reveals that the simple average of implied volatility performs better for American options than the simulated Heston implied volatility model. However, introducing neural networks significantly enhances the accuracy of implied volatility estimates.

Our results demonstrate that Feedforward Neural Networks consistently provide decent results in improving implied volatility for both models. However, we have discovered that the Long Short-Term Memory neural network correction yields even more stable results in predicting implied volatility and outperforms its benchmark, the FNN. This finding is significant as it demonstrates that the unique architecture of LSTM networks enables them to learn the temporal patterns of the data and make more accurate predictions. Interestingly, we have found that the Gated Recurrent Unit and Convolutional Neural Network are more sensitive to the size of the dataset, with worse results for smaller datasets but some of the best results for larger datasets. This finding suggests that these types of neural networks may require a certain level of data complexity to function optimally. The study indicates that the research question about the superior performance of time-based neural networks compared to FNNs has been confirmed.

However, to fully appreciate the implications of this conclusion, it is essential to consider the broader context of the research. One possible avenue for future research is to utilize a broader range of hyperparameters to tune neural networks. In this study, we only considered a limited set of hyperparameters, as LSTM, GRU, and CNN networks are computationally expensive. Furthermore, a larger number of neural networks can be used for non-parametric correction. As recent studies (citeTransformers) have suggested, the rising popularity of transformer neural networks can be employed as an additional source of time series-dependent neural networks to tune. As our results demonstrated the superior performance of Black Scholes-based correction for American-type options, further research could lead to the discovery of stochastic volatility models as PDE and non-PDE-based methods (**?**), which would outperform models under constant volatility for equity options with two-step approximation approach of .

Overall, our findings demonstrate the effectiveness of non-parametric correction of implied volatility models for American options, when combined with neural networks. The insights gained from this study can be applied in various financial settings, including option pricing, risk management, and portfolio optimization. Our findings provide practitioners with a powerful tool to enhance their decision-making processes and achieve superior financial outcomes.

References

Achdou, Y., Indragoby, G., and Pironneau, O. (2004). Volatility calibration with american options. *Methods and Applications of Analysis*, 11(4):533–556.

Akiba, T., Sano, S., Yanase, T., Ohta, T., and Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 2623–2631.

Almeida, C., Fan, J., Freire, G., and Tang, F. (2022). Can a machine correct option pricing models? *SSRN Electronic Journal*.

Andersen, T. G., Fusari, N., and Todorov, V. (2015). The risk premia embedded in index options. *Journal of Financial Economics*, 117(3):558–584.

Aruoba, S. B., Diebold, F. X., and Scotti, C. (2009). Real-time measurement of business conditions. *Journal of Business & Economic Statistics*, 27(4):417–427.

Audrino, F. and Colangelo, D. (2010). Semi-parametric forecasts of the implied volatility surface using regression trees. *Statistics and Computing*, 20:421–434.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The quarterly journal of economics*, 131(4):1593–1636.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–54.

Bloch, D. A. and Böök, A. (2020). Predicting future implied volatility surface using tdbp-learning. *DecisionSciRN: Decision-Making & Risk Management (Topic)*.

Bollerslev, T., Todorov, V., and Xu, L. (2015). Tail risk premia and return predictability. *Journal of Financial Economics*, 118(1):113–134.

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

Burkovska, O., Gass, M., Glau, K., Mahlstedt, M., Schoutens, W., and Wohlmuth, B. (2018). Calibration to american options: numerical investigation of the de-americanization method. *Quantitative finance*, 18(7):1091–1113.

Cao, J., Chen, J., and Hull, J. (2020). A neural network approach to understanding implied volatility movements. *Quantitative Finance*, 20(9):1405–1413.

Carr, P., Jarrow, R., and Myneni, R. (1992). Alternative characterizations of american put options. *Mathematical Finance*, 2(2):87–106.

Carr, P. and Wu, L. (2010). Stock options and credit default swaps: A joint framework for valuation and estimation. *Journal of Financial Econometrics*, 8(4):409–449.

Carr, P. and Wu, L. (2016). Analyzing volatility risk and risk premium in option contracts: A new theory. *Journal of Financial Economics*, 120(1):1–20.

Chen, S. and Zhang, Z. (2019). Forecasting implied volatility smile surface via deep learning and attention mechanism. *Available at SSRN 3508585*.

Chockalingam, A. and Muthuraman, K. (2011). American options under stochastic volatility. *Operations Research*, 59(4):793–809.

Chung, J., Gulcehre, C., Cho, K., and Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.

Clevert, D.-A., Unterthiner, T., and Hochreiter, S. (2015). Fast and accurate deep network learning by exponential linear units (elus). *arXiv preprint arXiv:1511.07289*.

Cont, R. and Da Fonseca, J. (2002). Dynamics of implied volatility surfaces. *Quantitative Finance*, 2(1):45–60.

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

Fodor, A., Stowe, D. L., and Stowe, J. D. (2017). Option implied dividends predict dividend cuts: Evidence from the financial crisis. *Journal of business finance & accounting*, 44(5-6):755–779.

Gatheral, J. (2004). A parsimonious arbitrage-free implied volatility parameterization with application to the valuation of volatility derivatives.

Gatheral, J. and Jacquier, A. (2014). Arbitrage-free svi volatility surfaces. *Quantitative Finance*, 14(1):59–71.

Gu, S., Kelly, B., and Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5):2223–2273.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–43.

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

Horvath, B., Muguruza, A., and Tomas, M. (2021). Deep learning volatility: a deep neural network perspective on pricing and calibration in (rough) volatility models. *Quantitative Finance*, 21(1):11–27.

Huang, Z., Yang, F., Xu, F., Song, X., and Tsui, K.-L. (2019). Convolutional gated recurrent unit–recurrent neural network for state-of-charge estimation of lithium-ion batteries. *Ieee Access*, 7:93139–93149.

Jacka, S. D. (1991). Optimal stopping and the american put. *Mathematical Finance*, 1(2):1–14.

Kim, I. J. (1990). The analytic valuation of american options. *The Review of Financial Studies*, 3(4):547–572.

Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kou, S. G. (2002). A jump-diffusion model for option pricing. *Management Science*, 48(8):1086–1101.

Lagnado, R., Osher, S., et al. (1997). A technique for calibrating derivative security pricing models: numerical solution of an inverse problem. *Journal of computational finance*, 1(1):13–25.

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4):541–551.

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. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1):125–144.

Muthuraman, K. (2008). A moving boundary approach to american option pricing. *Journal of Economic Dynamics and Control*, 32(11):3520–3537.

Ning, B., Jaimungal, S., Zhang, X., and Bergeron, M. (2021). Arbitrage-free implied volatility surface generation with variational autoencoders. *arXiv preprint arXiv:2108.04941*.

Ntakaris, A., Mirone, G., Kanniainen, J., Gabbouj, M., and Iosifidis, A. (2019). Feature engineering for mid-price prediction with deep learning. *Ieee Access*, 7:82390–82412.

Prechelt, L. (1998). Early stopping-but when? In *Neural Networks: Tricks of the trade*, pages 55–69. Springer.

Pérez-Enciso and Zingaretti, L. (2019). A guide for using deep learning for complex trait genomic prediction. *Genes*, 10:553.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386.

Shiraya, K. and Takahashi, A. (2017). Pricing average and spread options under local-stochastic volatility jump-diffusion models online appendix. (CARF-F-412).

Todorov, V. (2019). Nonparametric spot volatility from options. *Annals of Applied Probability*, 29:3590–3636.

Yan, D., Lin, S., Hu, Z., and Yang, B.-Z. (2022). Pricing american options with stochastic volatility and small nonlinear price impact: A pde approach. *Chaos, Solitons & Fractals*, 163:112581.

Zeng, Y. and Klabjan, D. (2019). Online adaptive machine learning based algorithm for implied volatility surface modeling. *Knowledge-Based Systems*, 163:376–391.

Zhang, W., Li, L., and Zhang, G. (2021). A two-step framework for arbitrage-free prediction of the implied volatility surface.

Appendix

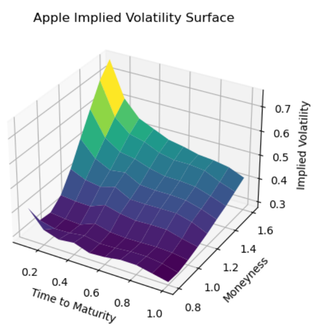

Figure 11: Implied volatility surface of Apple.

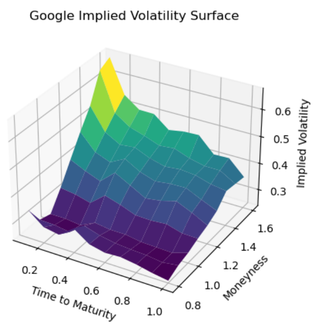

Figure 12: Implied volatility surface of Google.

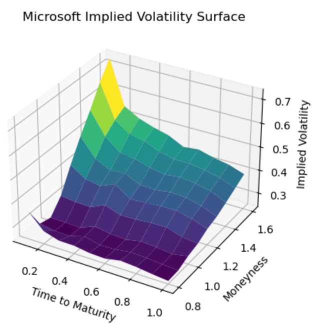

Figure 13: Implied volatility surface of Microsoft.

Table 6: Summary of Apple Stock Option Implied Volatility Statistics

Days to Maturity	**Moneyness**	**Number**	**Mean IV**	**Std. dev. IV**
Short	(0.8, 0.9]	2384	0.300	0.101
Short	(0.9, 0.97]	7402	0.246	0.083
Short	(0.97, 1.03]	4793	0.236	0.078
Short	(1.03, 1.1]	7150	0.271	0.096
Short	(1.1, 1.6]	5362	0.376	0.145
Long	(0.8, 0.9]	7000	0.251	0.072
Long	(0.9, 0.97]	8255	0.247	0.072
Long	(0.97, 1.03]	5049	0.237	0.075
Long	(1.03, 1.1]	7722	0.275	0.089
Long	(1.1, 1.6]	11378	0.337	0.117
Total	(0.8, 1.6]	66602	0.279	0.105

This table provides a summary of the Apple stock option implied volatility statistics for the period from December 31, 2018, to December 31, 2021. The table shows the count of options, the average implied volatility percentage, and the standard deviation of implied volatilities percentage for various time-to-maturity and moneyness categories. Short-term maturity options are considered as (20, 60] days to maturity, and long-term options as (60, 240], respectively.

Table 7: Summary of Google Stock Option Implied Volatility Statistics

Days to Maturity	**Moneyness**	**Number**	**Mean IV**	**Std. dev. IV**
Short	(0.8, 0.9]	9833.0	0.296	0.090
Short	(0.9, 0.97]	20952.0	0.259	0.082
Short	(0.97, 1.03]	14325.0	0.254	0.070
Short	(1.03, 1.1]	22383.0	0.296	0.089
Short	(1.1, 1.6]	24978.0	0.394	0.133
Long	(0.8, 0.9]	16321.0	0.287	0.093
Long	(0.9, 0.97]	25635.0	0.271	0.078
Long	(0.97, 1.03]	17156.0	0.261	0.076
Long	(1.03, 1.1]	25674.0	0.300	0.081
Long	(1.1, 1.6]	35866.0	0.372	0.117
Total	(0.8, 1.6]	213346.0	0.308	0.108

This table provides a summary of the Google stock option implied volatility statistics for the period from December 31, 2018, to December 31, 2021. The table shows the count of options, the average implied volatility percentage, and the standard deviation of implied volatilities percentage for various time-to-maturity and moneyness categories. Short-term maturity options are considered as (20, 60] days to maturity, and long-term options as (60, 240], respectively.

Table 8: Summary of Microsoft Stock Option Implied Volatility Statistics

Days to Maturity	**Moneyness**	**Number**	**Mean IV**	**Std. dev. IV**
Short	(0.8, 0.9]	5598.0	0.297	0.077
Short	(0.9, 0.97]	9139.0	0.249	0.080
Short	(0.97, 1.03]	4392.0	0.249	0.083
Short	(1.03, 1.1]	8717.0	0.287	0.088
Short	(1.1, 1.6]	13912.0	0.410	0.123
Long	(0.8, 0.9]	14683.0	0.267	0.070
Long	(0.9, 0.97]	13580.0	0.263	0.075
Long	(0.97, 1.03]	6339.0	0.261	0.078
Long	(1.03, 1.1]	11964.0	0.295	0.085
Long	(1.1, 1.6]	31152.0	0.369	0.098
Total	(0.8, 1.6]	119679.0	0.313	0.106

This table provides a summary of the Microsoft stock option implied volatility statistics for the period from December 31, 2018, to December 31, 2021. The table shows the count of options, the average implied volatility percentage, and the standard deviation of implied volatilities percentage for various time-to-maturity and moneyness categories. Short-term maturity options are considered as (20, 60] days to maturity, and long-term options as (60, 240], respectively.

Table 9: Summary of IBM Stock Option Implied Volatility Statistics

Days to Maturity	**Moneyness**	**Number**	**Mean IV**	**Std. dev. IV**
Short	(0.8, 0.9]	2384.0	0.299	0.101
Short	(0.9, 0.97]	7402.0	0.246	0.083
Short	(0.97, 1.03]	4793.0	0.236	0.078
Short	(1.03, 1.1]	7150.0	0.271	0.096
Short	(1.1, 1.6]	5362.0	0.376	0.145
Long	(0.8, 0.9]	7000.0	0.251	0.072
Long	(0.9, 0.97]	8255.0	0.247	0.072
Long	(0.97, 1.03]	5049.0	0.237	0.075
Long	(1.03, 1.1]	7722.0	0.275	0.089
Long	(1.1, 1.6]	11378.0	0.337	0.117
Total	(0.8, 1.6]	66602.0	0.279	0.105

This table provides a summary of the IBM stock option implied volatility statistics for the period from December 31, 2018, to December 31, 2021. The table shows the count of options, the average implied volatility percentage, and the standard deviation of implied volatilities percentage for various time-to-maturity and moneyness categories. Short-term maturity options are considered as (20, 60] days to maturity, and long-term options as (60, 240], respectively.

Table 10: Summary of Amazon Stock Option Implied Volatility Statistics

Model	**Hyperparameter**	**Values**
FNN	hidden layer	1, 2, 3, 4, 5
	learning rate	1e-1, 1e-2, 1e-3
	batch size	128
	dropout rate	0.1, 0.2, 0.3, 0.4, 0.5
	epoch	200, 300, 500
	penalty	0.001, 0.01, 0.1, 1
	lstm units	32, 64
LSTM	dense units	32, 64
	dropout rate	0.1, 0.2, 0.3, 0.4, 0.5
	learning rate	0.1, 0.2, 0.3, 0.4, 0.5
GRU	gru units	32, 64
	dense units	32, 64
	dropout rate	0.1, 0.2, 0.3, 0.4, 0.5
	learning rate	0.1, 0.2, 0.3, 0.4, 0.5
CNN	conv filters	32, 64
	kernel size	3, 4, 5
	dense units	32, 64
	dropout rate	0.0001, 0.001, 0.01, 0.1

The tables provides summary of hyperparameters used to tune the FNN, LSTM, GRU and CNN neural network for pricing correction exercise.

Table 11: Black Scholes and Heston model corrected performance for Apple stock options

	Black Scholes					**Heston**				
	No NN	**FNN**	**LSTM**	**GRU**	**CNN**	**No NN**	**FNN**	**LSTM**	**GRU**	**CNN**
All	0.089	**0.066**	0.067	0.068	**0.066**	0.119	0.097	0.069	0.064	**0.061**
DOTMC	0.066	0.066	0.052	0.047	**0.043**	0.121	0.034	**0.251**	0.485	0.478
OTMC	0.067	0.068	0.055	**0.052**	0.054	1.939	1.938	1.739	**0.052**	0.054
ATM	0.073	0.074	0.063	0.064	**0.060**	2.441	**0.157**	0.412	2.315	2.214
OTMP	0.068	0.138	0.064	0.066	**0.061**	0.334	0.390	0.252	0.204	**0.195**
DOTMP	0.088	0.072	0.071	0.070	**0.066**	0.379	0.285	0.255	0.225	**0.185**
Short	0.101	0.091	0.075	0.072	**0.071**	1.919	**1.085**	1.252	1.548	1.344
Long	0.082	0.078	0.065	0.069	**0.061**	0.232	0.164	0.131	0.122	**0.104**

The table provides a summary of the pricing performance of the FNN, LSTM, GRU, and 1D CNN neural networks combined with Black Scholes and the Heston model measured in RMSE (root mean squared error) per category.

Table 12: Black Scholes and Heston model corrected performance for Google stock options

	Black Scholes					Heston				
	No NN	**FNN**	**LSTM**	**GRU**	**CNN**	**No NN**	**FNN**	**LSTM**	**GRU**	**CNN**
All	0.100	0.084	0.060	0.060	**0.056**	0.182	0.107	**0.081**	0.121	0.122
DOTMC	0.085	0.085	0.051	**0.047**	0.051	0.374	0.296	0.206	0.205	**0.180**
OTMC	0.066	0.066	**0.045**	0.047	0.046	1.555	**0.177**	0.270	1.315	1.304
ATM	0.059	0.062	**0.044**	0.208	**0.044**	2.429	**0.186**	0.234	2.295	2.301
OTMP	0.072	0.074	0.051	**0.049**	0.051	0.320	**0.036**	0.235	0.191	0.153
DOTMP	0.116	0.111	0.069	0.073	**0.065**	0.219	0.141	0.117	0.109	**0.103**
Short	0.102	0.092	**0.052**	0.057	**0.052**	**0.065**	0.481	0.313	0.456	0.412
Long	0.098	0.087	0.061	**0.057**	0.061	1.267	0.588	0.510	**0.298**	0.777

The table provides a summary of the pricing performance of the FNN, LSTM, GRU, and 1D CNN neural networks combined with Black Scholes and the Heston model measured in RMSE (root mean squared error) per category.

Table 13: Black Scholes and Heston model corrected performance for Microsoft stock options

	Black Scholes					Heston				
	No NN	**FNN**	**LSTM**	**GRU**	**CNN**	**No NN**	**FNN**	**LSTM**	**GRU**	**CNN**
All	0.087	0.062	0.059	0.127	0.122	0.199	0.091	0.071	0.065	0.122
DOTMC	0.046	0.049	0.057	0.078	0.055	1.694	0.114	0.108	1.513	1.474
OTMC	0.046	0.070	0.049	0.066	0.047	1.957	0.196	0.076	0.564	0.689
ATM	0.049	0.054	0.048	0.049	0.044	2.161	0.098	0.0657	0.066	0.314
OTMP	0.050	4.410	0.049	0.049	0.049	0.325	0.211	0.098	0.154	0.278
DOTMP	0.082	0.062	0.065	0.079	0.062	1.286	0.314	0.096	0.063	0.065
Short	0.102	0.054	0.069	0.079	0.061	1.885	0.110	0.124	0.091	0.316
Long	0.079	0.068	0.062	0.088	0.070	0.336	0.564	0.478	0.312	0.890

The table provides a summary of the pricing performance of the FNN, LSTM, GRU, and 1D CNN neural networks combined with Black Scholes and the Heston model measured in RMSE (root mean squared error) per category.

Table 14: Black Scholes and Heston model corrected performance for IBM stock options

	Black Scholes					**Heston**				
	No NN	**FNN**	**LSTM**	**GRU**	**CNN**	**No NN**	**FNN**	**LSTM**	**GRU**	**CNN**
All	0.073	0.067	**0.059**	0.127	0.122	0.199	0.091	0.071	**0.065**	0.122
DOTMC	0.055	0.054	**0.039**	0.131	0.111	0.277	0.241	0.165	0.162	**0.111**
OTMC	0.044	0.050	**0.038**	1.846	1.963	2.216	0.787	**0.314**	0.341	1.963
ATM	0.047	0.056	**0.042**	1.829	1.852	2.128	**0.190**	0.457	0.338	1.852
OTMP	0.060	0.062	**0.051**	0.287	0.116	0.336	0.385	0.249	0.351	**0.116**
DOTMP	0.090	0.102	**0.075**	0.324	0.297	0.742	0.562	0.388	0.353	**0.297**
Short	0.075	0.077	**0.060**	1.775	1.521	2.138	0.512	0.952	**0.500**	1.521
Long	0.071	0.079	0.054	**0.044**	0.061	0.211	0.209	0.183	0.187	**0.061**

The table provides a summary of the pricing performance of the FNN, LSTM, GRU, and 1D CNN neural networks combined with Black Scholes and the Heston model measured in RMSE (root mean squared error) per category.

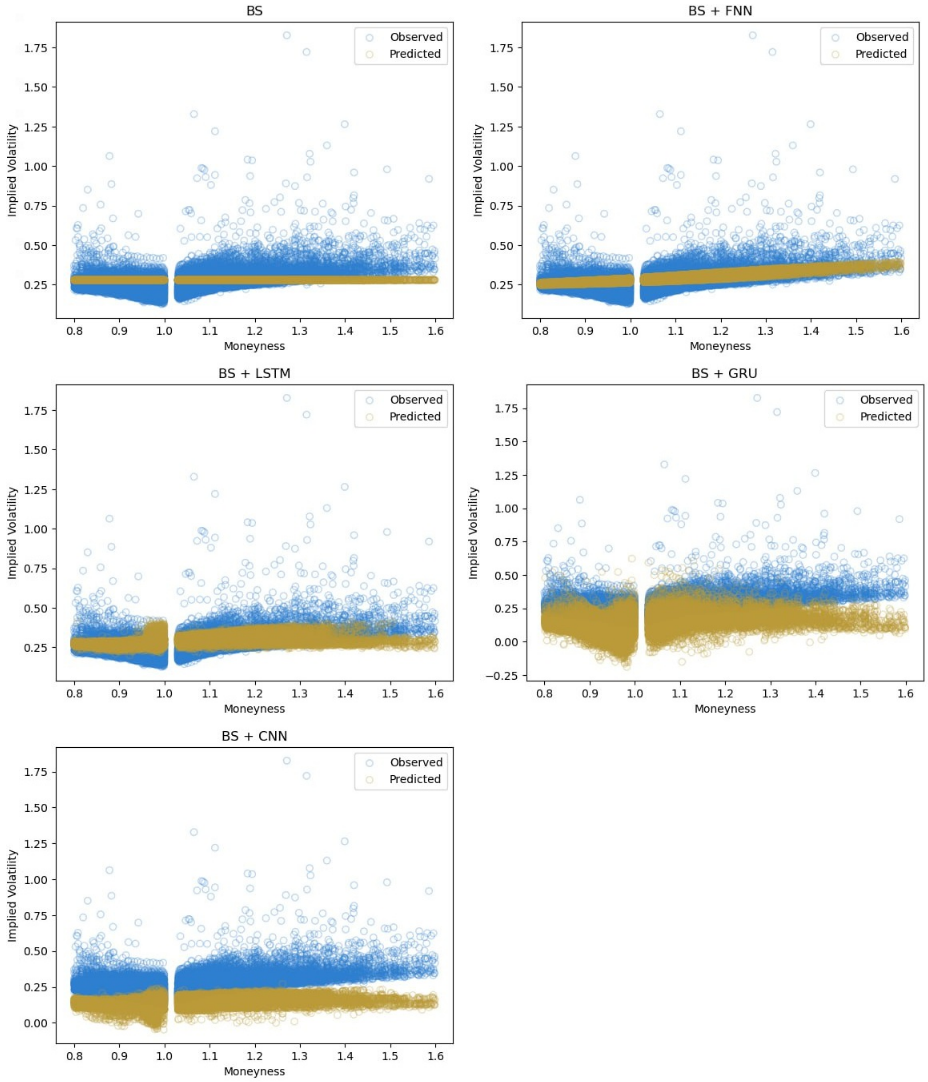

Figure 14: Prediction in the Apple option panel from 31 December 2020 to 31 December 2021.

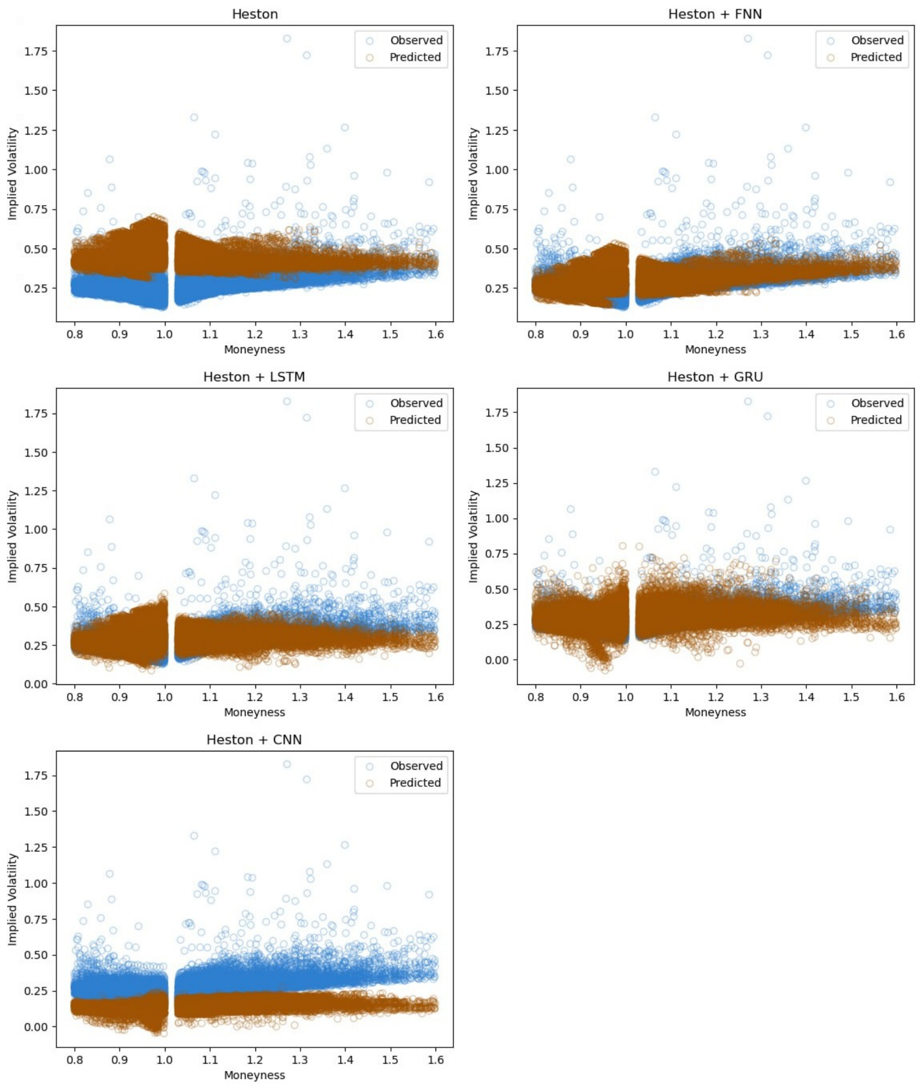

Figure 15: Prediction in the Apple option panel from 31 December 2020 to 31 December 2021.

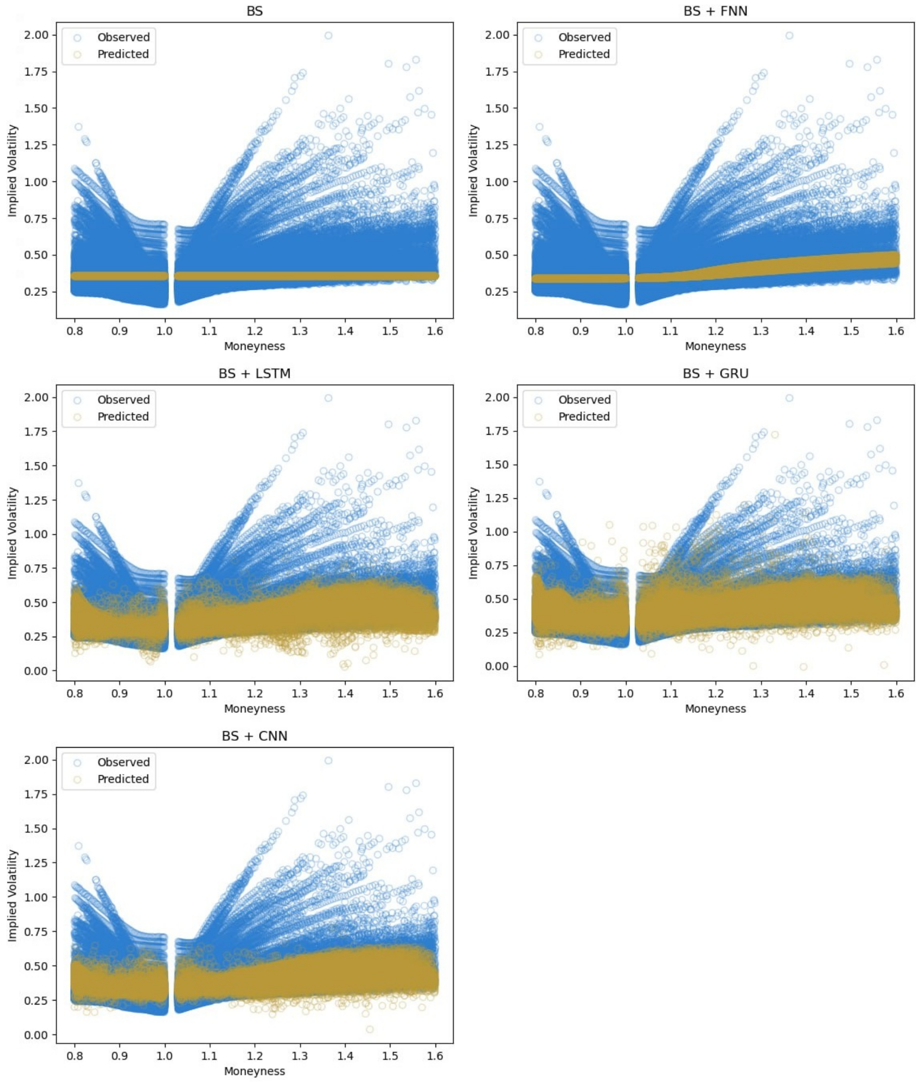

Figure 16: Prediction in the Amazon option panel from 31 December 2020 to 31 December 2021.

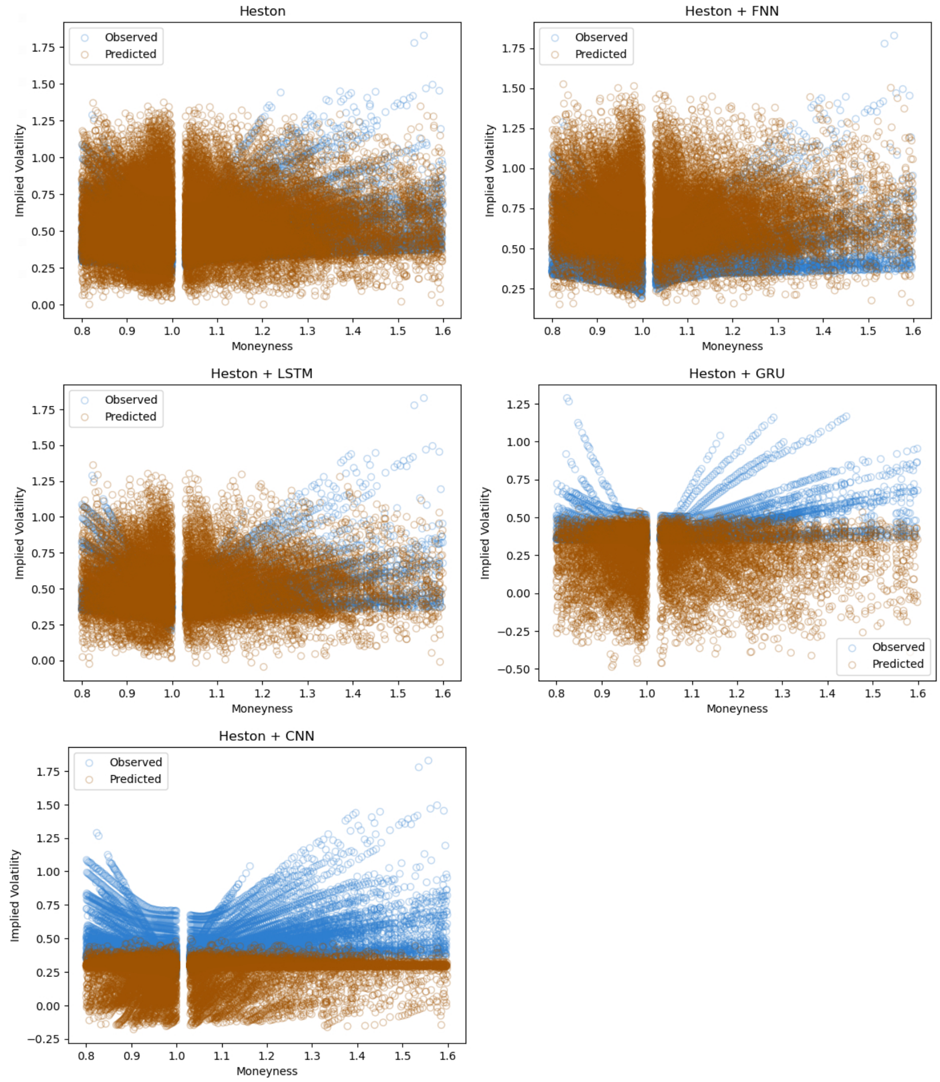

Figure 17: Prediction in the Amazon option panel from 31 December 2020 to 31 December 2021.

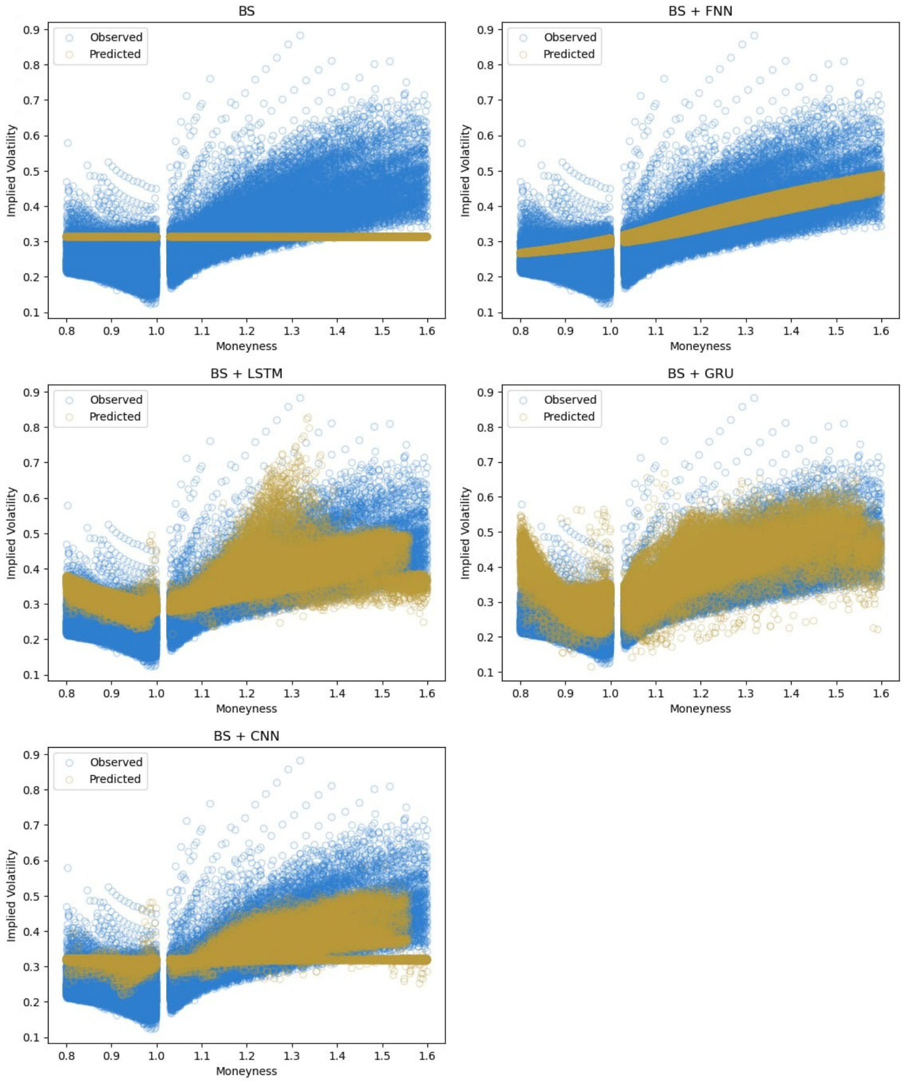

Figure 18: Prediction in the Microsoft option panel from 31 December 2020 to 31 December 2021.

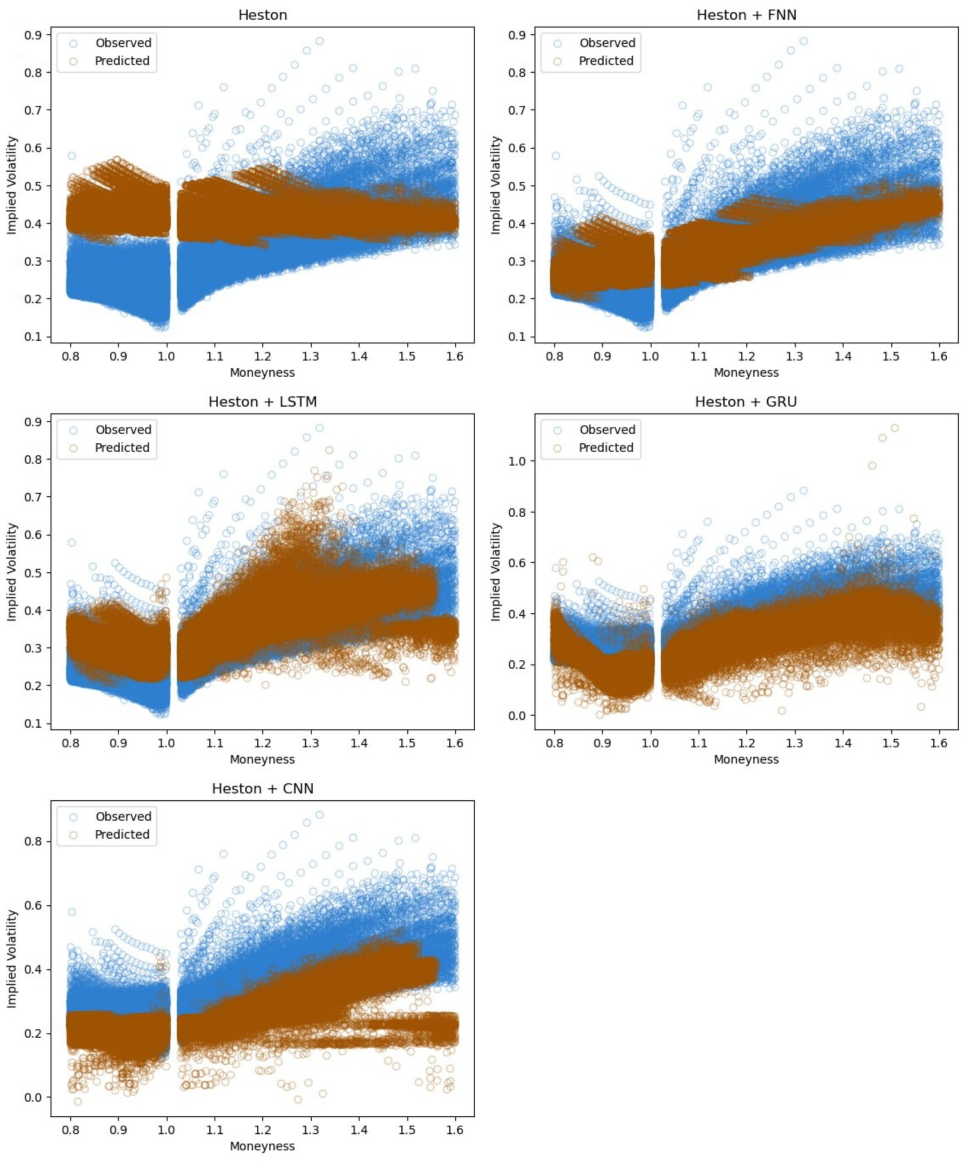

Figure 19: Prediction in the Microsoft option panel from 31 December 2020 to 31 December 2021.

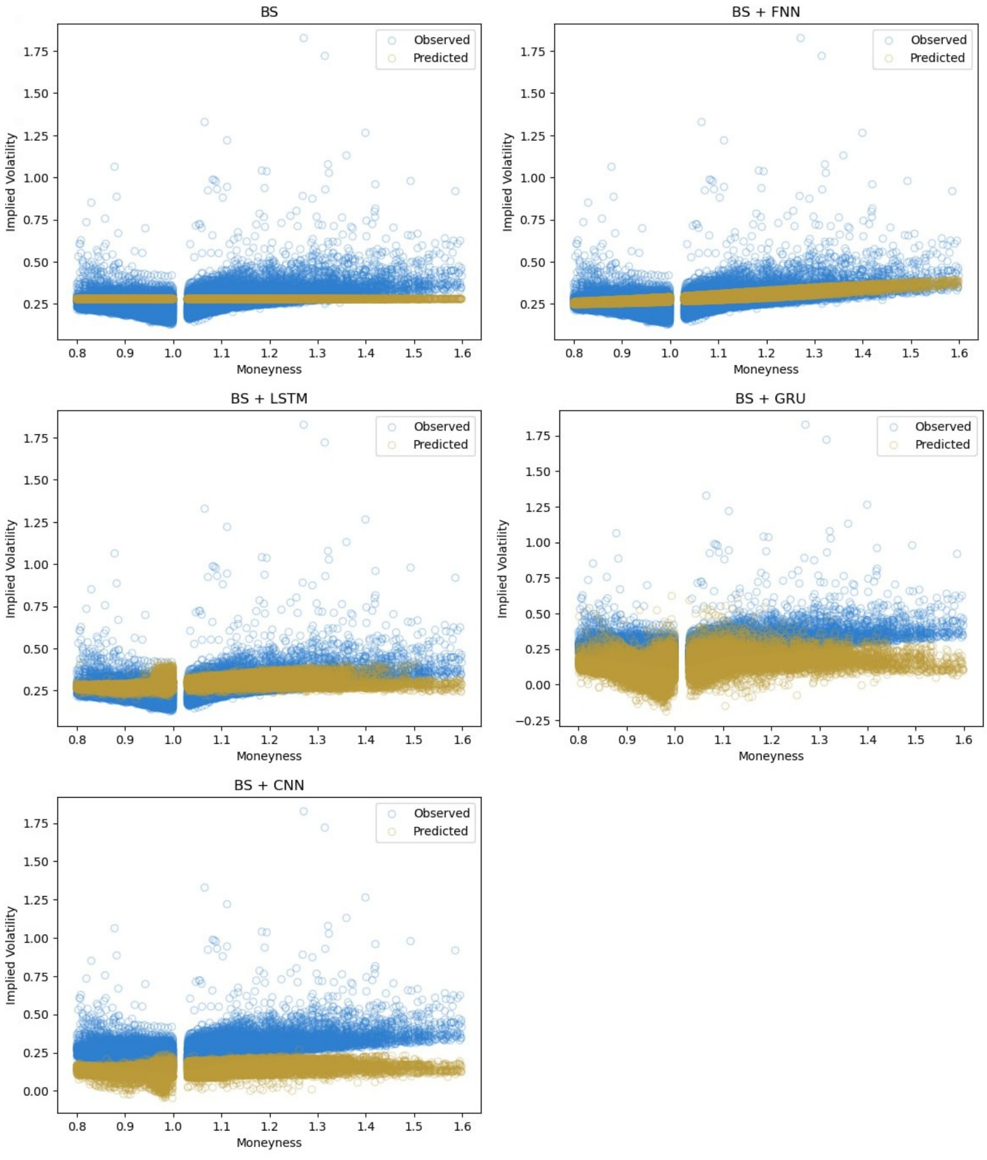

Figure 20: Prediction in the IBM option panel from 31 December 2020 to 31 December 2021.

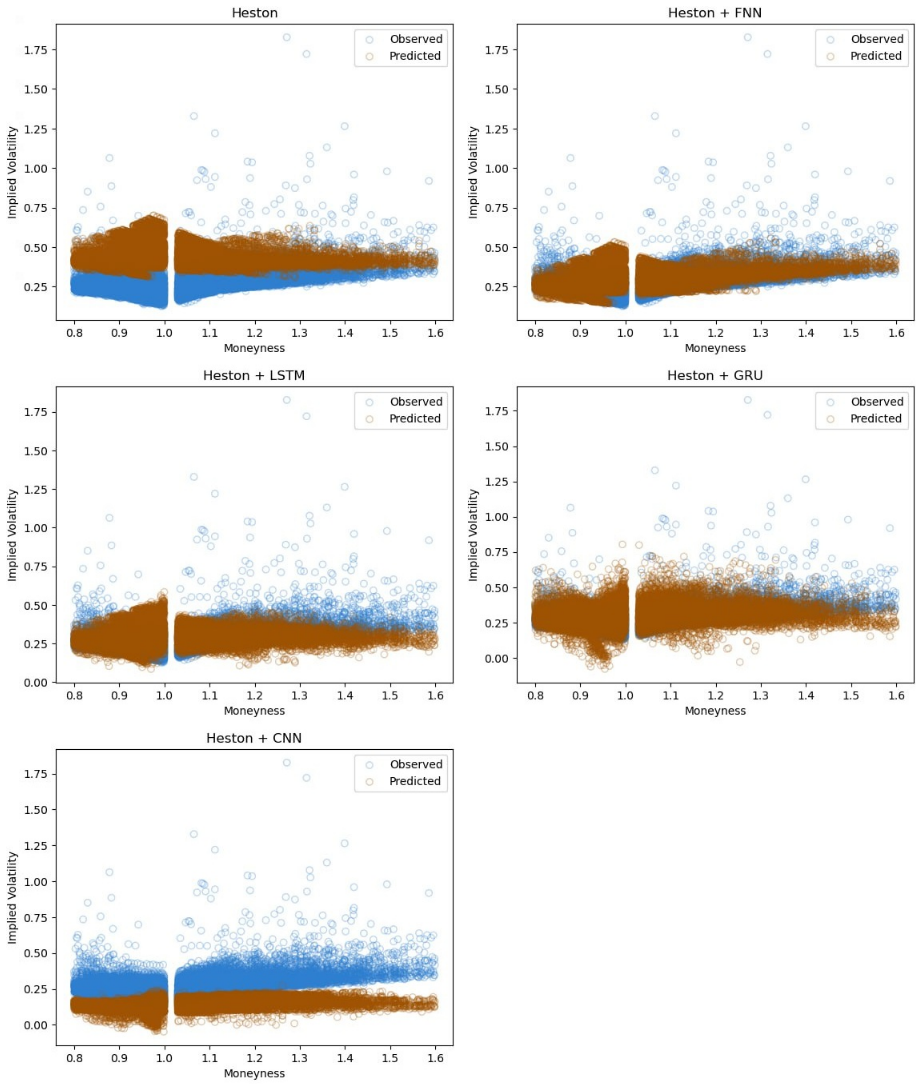

Figure 21: Prediction in the IBM option panel from 31 December 2020 to 31 December 2021.