Leung *et al. Financial Innovation (2024) 10:130*
https://doi.org/10.1186/s40854-024-00631-5

Financial Innovation

RESEARCH

Open Access

Deterministic modelling of implied volatility in cryptocurrency options with underlying multiple resolution momentum indicator and non-linear machine learning regression algorithm

F. Leung[1*^], M. Law[2*] and S. K. Djeng[2]

^F. Leung is deceased.

*Correspondence: leungfranklin33@gmail.com; ynbugy_law@yahoo.com

[1] Department of Supply Chain and Information Management, The Hang Seng University of Hong Kong, HKSAR, China
[2] Proton Therapy Pte Ltd, 1 Biopolis Drive, Biopolis, Singapore

Abstract

Modeling implied volatility (IV) is important for option pricing, hedging, and risk management. Previous studies of deterministic implied volatility functions (DIVFs) propose two parameters, moneyness and time to maturity, to estimate implied volatility. Recent DIVF models have included factors such as a moving average ratio and relative bid-ask spread but fail to enhance modeling accuracy. The current study offers a generalized DIVF model by including a momentum indicator for the underlying asset using a relative strength index (RSI) covering multiple time resolutions as a factor, as momentum is often used by investors and speculators in their trading decisions, and in contrast to volatility, RSI can distinguish between bull and bear markets. To the best of our knowledge, prior studies have not included RSI as a predictive factor in modeling IV. Instead of using a simple linear regression as in previous studies, we use a machine learning regression algorithm, namely random forest, to model a nonlinear IV. Previous studies apply DVIF modeling to options on traditional financial assets, such as stock and foreign exchange markets. Here, we study options on the largest cryptocurrency, Bitcoin, which poses greater modeling challenges due to its extreme volatility and the fact that it is not as well studied as traditional financial assets. Recent Bitcoin option chain data were collected from a leading cryptocurrency option exchange over a four-month period for model development and validation. Our dataset includes short-maturity options with expiry in less than six days, as well as a full range of moneyness, both of which are often excluded in existing studies as prices for options with these characteristics are often highly volatile and pose challenges to model building. Our in-sample and out-sample results indicate that including our proposed momentum indicator significantly enhances the model's accuracy in pricing options. The nonlinear machine learning random forest algorithm also performed better than a simple linear regression. Compared to prevailing option pricing models that employ stochastic variables, our DIVF model does not include stochastic factors but exhibits reasonably good performance. It is also easy to compute due to the availability of real-time RSIs. Our findings indicate our enhanced DIVF model offers significant improvements and may be an excellent alternative to existing option pricing models that are primarily stochastic in nature.

Keywords: Implied volatility, Cryptocurrency options, Momentum indicator, Relative strength index, Machine learning, Random Forest regression, Black–Scholes-Merton equation

Introduction

The Black–Scholes-Merton (BSM) option pricing model (Black & Scholes 1973) estimates the prices of derivatives based on observable prices of other securities while considering the impact of time and other risk factors. Developed in 1973, it is still considered one of the most popular approaches to pricing options contracts. The model assumes the underlying asset's price follows a geometric Brownian motion with a lognormal distribution and constant volatility. Using the historical volatility of the underlying asset and other parameters such as interest rates as inputs, the BSM model calculates option prices. However, some criticize the model's assumption that asset prices are log normally distributed because stock returns exhibit both fat-tailed marginal distributions and volatility clustering (Fama 1965; Mandelbort 1966).

Implied volatility (IV) is calculated as the value that fits an option model's price to the market price. IV is interpreted as the expected volatility implied by the market, while historical volatility reflects the past trading range of an underlying asset. Retail option traders may be more interested in IV due to its forward-looking nature. It estimates an underlying asset's future price movement according to changes in the price of a corresponding option and can be used as a useful approach in comparing prices for different underlying assets, strike prices, and maturities. Traders may quote option prices in terms of implied volatilities, and options exchanges usually provide implied volatilities in real time. A three-dimensional graph known as an implied volatility surface (IVS) shows the implied volatilities of an asset's options across a range of strike prices and expiration dates (Daglish et al. 2007; Cont & Fonseca 2002). Figure 1 is an example of an IVS (Wang et al. 2017). Instead of strike prices, moneyness, i.e., the ratio between the current asset price (S) and the strike price (K), is often used to construct an IVS analysis. The IVS should, in theory, be flat, but in reality it is curved due to market variations

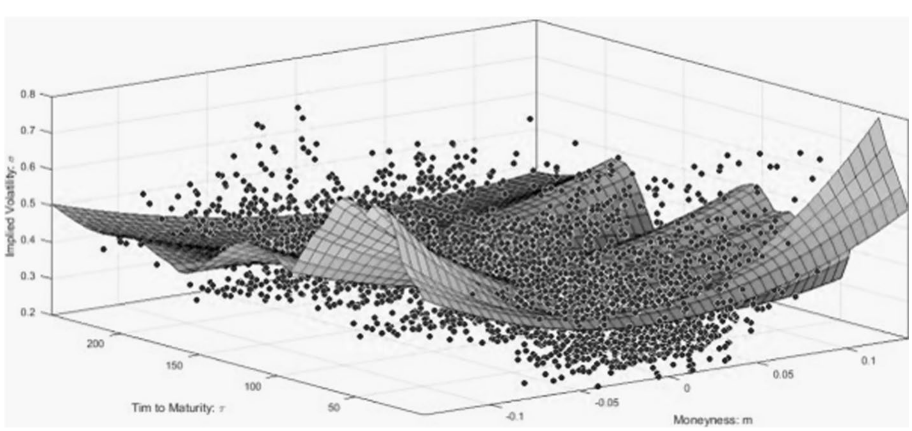

Fig. 1 The implicit volatility surface for option chains over a period of time (Wang 2017)

(Fig. 1). As every option chain is different, the shape of a volatility surface may be wavy across moneyness and time to expiry (Fengler, Hardle & Mammen 2005; Cont and Fontseca, 2002). When an IVS is constructed for certain time periods, multiple IV values are observed for the same moneyness and time to expiry, implying that implied volatilities depend on market volatility at different points in time.

Two lines of research suggest expansions to the BSM model: jump-diffusion models featuring a Poisson-driven jump process, and stochastic volatility models. However, as the stochastic parameters are estimated based on the past behavior of underlying asset prices over a specific timeframe, their estimates may only be applicable for that time period or under similar market conditions. This means that if parameter estimates are derived from data during a relatively unstable period, option prices calculated using the model in a calmer market may be overestimated.

Modeling implied volatility (IV) directly from option prices observed in the market is another approach to option pricing. By successfully modeling IV, it can be used as a parameter in the BSM equation to calculate more accurate options prices than those obtained from BSM using realized volatilities. Early studies on deterministic IV functions only included moneyness and time to expiry (Dumas et. al. 1998; henceforth DFW), but later attempts to include additional factors such as a moving average ratio and relative bid-ask spread have failed to improve model accuracy (Pena et al. 1999; 2001). This study proposes using an indicator of an asset's underlying momentum, known as a relative strength index (RSI), with multiple time resolutions as a factor in modeling IV because momentum is often considered by investors and speculators in their trading decisions. To our knowledge, no prior studies have included RSI as a predictive factor in IV modeling. Unlike volatility, RSI can differentiate between bullish and bearish markets and is commonly used by traders to identify overbought or oversold situations.

In this study, we use a machine learning regression algorithm known as random forest to model the non-linear IV surface instead of the generalized linear regression approach used in previous studies. While past studies focus on traditional financial assets such as stocks, stock indices and foreign exchange indices, this study examines options on Bitcoin, the largest cryptocurrency, which poses significant challenges due to its high volatility. Unlike options on traditional financial assets, Bitcoin options have not been extensively studied because Deribit – the leading options exchange for cryptocurrencies – was only launched in 2018 which means data may be available for four only years. Separately, Bitcoin options are more heavily traded than options on other underlying assets as they are traded continuously over a 24-h day, which means that more data points can be collected compared to other securities and foreign exchange rates. A nonlinear regression analysis using a machine learning algorithm outperforms the traditional linear regression analysis used in past studies for model development and validation.

The research gaps identified in this study are as follows:

1. Prior studies have not succeeded in improving the accuracy of the DIVF model proposed in DFW. Including the RSI momentum indicator has not been done before and we show it can significantly enhance the accuracy of a DIVF model in modeling IV. Despite being less researched over the past two decades, DIVF remains relevant

for option pricing due to its simple formulation and ease of computation compared to other stochastic option pricing models.
2. This study addresses near to at-the-money and close-to-expiration options, which have been excluded from previous research due to their large price fluctuations during market stress that presents challenges to modeling. To achieve better IV modeling, we use a nonlinear machine learning regression algorithm in place of the linear regression algorithms used in previous DIVF studies.
3. This study focuses on Bitcoin options data, which has received relatively less attention than more traditional financial assets such as listed securities and foreign exchange. The highly volatile nature of Bitcoin markets poses additional challenges to model-building but our model demonstrates good performance even under these conditions.

Literature review

Option pricing models

Option pricing models have a significant impact on the trading, investing and corporate finance. Chew and Stewart (2022) review the practical significance of option pricing models using case studies. Professional options traders can be seen walking the trading floors armed with commercial versions of various option pricing models. Option pricing theory also has an important impact on corporate finance practitioners. Option pricing methods are the most promising way of attempting to quantify the value of merger and acquisition possibilities as a "portfolio of options" i.e., a firm's option to invest in second-stage, third-stage, or even later-stage projects. The application of option pricing methods contributes to the search for a firm's optimal capital structure.

Table 1 offers a high-level overview of option pricing models with their pros and cons. Rubinstein (1985, 1994), among others, provides evidence that implied volatilities tend to be higher for options that are deeply in- or out-of-the-money compared to at-the-money options. Bates (2000) shows that relative to call options, put options are underpriced by the BSM formula, which, in turn, suggests that the IV curve is downward-sloping in the strike price. For example, put options with lower moneyness tend to have higher implied volatilities than those with higher moneyness. The volatility surface is often observed to have an inverted volatility smile, where options with a short time to maturity have volatilities that are multiples of volatilities for options with longer maturities, particularly in periods of high market stress (Bakshi et al. 1997). Kim (2009) finds that in the equity options market the IV of deep out-of-the-money puts is larger than that of deep out-of-the-money calls, producing a volatility "smirk," or a volatility smile that is skewed in one direction. This indicates that out-of-the-money puts have higher valuations than comparable out-of-the-money calls. Many studies show that observed option price behavior deviates from BSM estimates, especially during periods of market stress. For example, Voukelatos & Verousis (2019) find that in the US equity market investors tend to herd during market stress and the options market is characterized by a higher level of IV and more negative implied skewness. Forlicz (2011) notes that in Poland's stock market, call option prices quoted in a bull market are underestimated relative to model prices during a bear market. Generally, in bull and bear markets, put options are overestimated relative to prices resulting from the BSM model.

Table 1 High level comparison between option pricing models

Approach: Historical Volatility/ Implied Volatility-Deterministic/ Implied Volatility-Dynamic	**References**	**Pros**	**Cons**
Historical Volatility	BSM (Black & Scholes 1973)	Closed Form Solution	Assuming return log-normal distribution and constant volatility; Comparatively lower accuracy both for in-sample and out-sample datasets against later models
	Pure jump model (Cox & Ross 1976), Jump-diffusion model (Merton 1976) Lévy process model (Carr & Wu 2004; Feng et. al. 2020) Variance Gamma Model (Madan & Seneta 1990; Madan et al. 1998; Mehrdoust & Samimi 2020)	Ability to capture the effects of jumps and discontinuities of asset price, account for skewness and kurtosis in asset return, later models can cater for asymmetric nature of volatility (distinguishing up and down moves)	Closed form solution not available and solution has to be obtained through Monte Carlo simulation or Fast Fourier Transform. The model parameters are estimated from historical dataset and may not be stable especially in volatility regime switching
	Stochastic Volatility Model (Heston 1993; Christoffersen et. al. 2009; Camera & Heston 2008; Mehrdoust et al. 2021) GARCH Model (Heston & Nandi 2000; Arunsingkarat 2021; Venter et al. 2020)	Closed Form solution, Later models introduce multi-factor stochastic models and models distinguishing ups and downs movement. GARCH model incorporates the observed path of historical spot prices to enhance the model accuracy	The model parameters are estimated from historical dataset and may not be stable especially in volatility regime switching
	Regime-switching stochastic volatility model (Elliot et. al. 2007; Goutte et. al. 2017; Siu & Elliott (2021), Escobar-Anel (2021)	Incorporates a Markov regime switching model to achieve a better in-sample matching with market IV	Closed form solution not available and solution has to be obtained via Monte Carlo simulation. Out-sample performance has not been illustrated
Implied Volatility-Deterministic	Original DIVF (DFW 1998)	Deterministic, simple linear model on two factors: moneyness and time to maturity	Prediction error remains large, not stable in pricing out-sample options even for a short time period ahead
	Ad-hoc BSM (DFW 1998)	Deterministic, simple linear model on two factors: moneyness and time to maturity, stable out-sample performance and used as benchmark against other optional pricing models	Function has to re-estimated every week to achieve good performance in out-sample
	Pena et. al. (1999, 2001)		Adding relative momentum(moving average ratio) and bid-ask spread failed to enhance model accuracy
Implied Volatility-Dynamic	Guo (2000) Cont & Fonseca (2002) Panigirtzoglous & Skiadopoulos(2004) Neumann & Skiadopoulos (2012) Guo, Han & Zhao (2014) Guo, Han & Lin (2017) Bloch (2020) Chen et al (2022)	Allows for greater flexibility in model parameters with long term and short term components, greater accuracy in modelling asset price dynamics, can predict the entire IVS for next time period (next month/week/day ahead)	It cannot fully address the time variation of IVS. The coefficients need to be re-estimated before prediction for next time periods (say next month, week or day ahead) for enhancing out-sample accuracy

To capture difference between market-observed option prices and prices produced by the BSM equation, one early approach was to construct implied binomial trees (Rubinstein 1994). Jackwerth and Rubinstein (1995) constructed a binomial or trinomial numerical procedure to achieve a perfect fit with observed option prices. Later, DFW conducted empirical tests of implied binomial trees and found that their pricing and hedging (out-of-sample) performance was even worse than that of their proposed ad hoc BSM model that used variable implied volatilities, and the stability of the volatility function was seen as the key reason.

Another approach is to fit generalized stochastic processes to option prices. Many such option pricing models have been proposed including the pure jump model in Cox and Ross (1976), the jump-diffusion model in Merton (1976), the stochastic volatility model in Heston (1993), the Lévy process model (Carr and Wu 2004; Feng et al. 2020) and the Variance Gamma model (Madan & Seneta 1990; Mehrdoust & Samimi 2020). These models extend the BSM theory to other mathematical methodologies resulting in more complicated formulas. Some of these models further differentiate up and down moves (Madan et al. 1998; Camera & Heston 2008) to address the asymmetric nature of volatility. As these models are stochastic, options are often priced using a Monte Carlo simulation. For certain models, the Fast Fourier Transform (FFT) (Carr & Madan 1999) can be used to compute option prices efficiently but FFT has its limitations (Nzokem 2021). Nevertheless, Heston (1993) derives a closed-form solution for a stochastic volatility model. Heston and Nandi (2000) further develop a closed-form formula for their generalized autoregressive conditional heteroskedasticity (GARCH) models in which option value is described as a function of the option price and the observed path of historical spot prices. GARCH models have been applied in modeling option prices for the Stock Exchange of Thailand (Arunsingkarat et al. 2021) and S&P 500 Volatility Index (VIX) option pricing (Venter et al. 2020). Regime-switching stochastic volatility models (Elliot et al. 2007; Goutte et al. 2017; Escobar-Anel et al. 2021; Siu & Elliott 2021) further incorporate a Markov regime-switching model to achieve a better in-sample match with observed option prices in a market with volatility shifts. Nevertheless, the importance of out-sample testing was pointed out by studies of empirical tests of various models (Bakshi et al. 1997; Kim 2009; Heston & Nandi 2000), as in-sample fit does not necessarily lead to out-sample fit. The parameters for some models must be reestimated frequently to achieve an acceptable out-of-sample fit.

Modeling implied volatility—deterministic approach

Another approach to option pricing is to model the IVS. If the IV function is accurate, the parameter can be fed into the BSM formula to calculate the option's estimated price. This should be a better estimate than what the BSM formula itself calculates, as the option's price behavior in the market was taken into consideration in modeling the IV. The DIVF is one such approach. Using another technique, the time-varying volatility of the FTSE 100 index has been estimated using the ordinary least square regression on time to maturity and moneyness (Ncube 1996). The model depends on strikes and the first and second powers of the time to expiry. Similarly, DFW (1998) propose several volatility models based on both strike and time to expiration and fit such models to S&P 500 index options over 5 year period. Their results outperform the implied binomial trees in Rubinstein (1994) in out-of-sample validation tests of S&P 500 index options but the prediction errors are still relatively large. By evaluating "how well each week's estimated volatility function values the same options one week later," DFW conclude that such static and deterministic volatility functions are unstable.

DFW develop an ad-hoc BSM model in which the volatility function is reestimated every week and the implied volatility obtained is used to calculate option prices. This ad-hoc BSM outperforms the originally proposed deterministic volatility functions and has since been used as a benchmark against other option pricing models as it has been proven to be difficult to beat in out-of-sample tests (Goncalves & Guidolin 2006). Rosenberg (2000) proposes a refined model that employs maturity-adjusted-proportional-moneyness (moneyness divided by the square root of time to expiry) and at-the-money IV. Alentorn (2004) tests models in DFW (1998) by replacing the function of the strike price with maturity-adjusted-proportional-moneyness. Pena et al. (1999) find that the relative momentum of the market (defined as the ratio of the current level against the three- month moving average) is weakly related to the degree of curvature of the volatility smile. Pena et al. (2001) estimate the volatility function using relative bid-ask spread, time to expiry, and exercise price, but the model performed even worse than the original BSM model.

Modeling implied volatility—dynamic approach

To address the time varying nature of IV, a dynamic approach has been proposed in which stochastic state variables driving changes in individual implied volatilities are identified and modeled. This dynamic approach has dominated most studies on IV function over the past two decades. Initially, some of these studies focused either on the term structure of only at-the-money implied volatilities (Rosenberg 2000) or separately analyzed IVS for different maturities (Cont & Fonseca 2002). Some of these earlier studies focus on the cross-section of an IV model but ignore the time dimension, while others model the time-series property of an arbitrarily chosen point on an IVS computed as the volatility implicit in contracts with a given moneyness and/or time-to-expiration (Harvey & Whaley 1992; Guo 2000; Brooks & Oozeer 2002). Later studies cover longer-term maturities and a wider range of moneyness. These studies usually reduce dimensionality via a principal component analysis by projecting the IVS onto some "latent" factors that explain the evolution of implied distributions through time (Christoffersen et al. 2009). Dynamic modeling is then conducted using the time series of the extracted principal components (Panigirtzoglous & Skiadopoulos 2004) or higher-order moments (Neumann and Skiadopoulos 2012). Christoffersen et al. (2008) propose component volatility models and decompose stochastic volatility into long- and short-term components. Guo et al. (2018) also find that short-maturity implied volatilities are more related to a short-term variance factor, while long-maturity implied volatilities are more related to long-term variance. More recent studies further investigate volatilities for different time horizons (short-, medium-, and long-term), exploring the explanatory variables driving the "latent" factors obtained from the principal component analysis. Bloch & Book (2021) suggest the explanatory variables are the spot price of the stock, time to maturity, trading volume, estimated volatility, and financial indicators such as the VIX. Chen et al. (2022) find that firm fundamentals have explanatory power regarding the shape of the IV curve for options on that firm's stock. Guo & Han (2014) propose three components: a long-term component driven by macroeconomic variables, a medium-term component driven by default risk, and a short-term component driven by financial market conditions. Models using the dynamic approach usually predict an IVS over a short period, from one day (Gonclaves & Guidon 2006; Guo & Han 2014; Wang et al. 2017) to one week. Guo et al. (2018) show that an IVS for at-the-money options is only useful for an out-of-sample forecast of implied volatility up to one week ahead for call options, and up to 20 days ahead for puts.

Cryptocurrencies

Cryptocurrencies have attracted increasing attention ever since Satoshi Nakamoto introduced Bitcoin, a peer-to-peer electronic transaction system (Nakamoto 2008). Bitcoin can be viewed as a decentralized digital currency that eliminates the need for central authorities such as banks or governments. As it can be traded on cryptocurrency exchanges, its use as an alternative financial asset is also studied extensively. Barson et al. (2022) note that the relationship between gold and some cryptocurrencies was more significant over longer periods during the COVID-19 pandemic than in the medium-term, and find a high persistence in the hedging properties of gold with Bitcoin. Asafo-Adjei et al. (2021) find there was a significant negative information flow from global equities to cryptocurrencies during the COVID-19 pandemic; hence, cryptocurrencies may be seen as a safe haven for global equities during uncertainty. However, Jalal et al. (2021) point out that, despite Bitcoin's weak correlation with traditional asset classes, Bitcoin's ability to function as a safe-haven currency and portfolio diversifier is debatable due to its lack of a regulatory framework and poor security. Harb et al. (2022) show that the cryptocurrency market is uncorrelated to the US stock market but not the US bond market. Bitcoin has been characterized by dramatic upward and downward price movements associated with high transaction volumes. There are numerous studies on modeling the volatility of cryptocurrencies (Chi & Hao 2020). Agyei et al. (2022) find that interdependencies between cryptocurrencies and a cryptocurrency implied volatility index (VCRIX) are high and mostly positive across investment horizons. Bouri et al. (2018) show evidence of long memory in Bitcoin price volatility, implying a Bitcoin market inefficiency that market participants and analysts may be able to exploit. Balcilar et al. (2017) reveal that Bitcoin trading volume could predict returns only when the market is not volatile and trading volume cannot predict Bitcoin volatility. Other cryptocurrencies' implied volatilities and volatility modeling are less studied. Agyei et al. (2022) conduct a multi-scale and time–frequency analysis of the degree of integration and lead-lag relationship among six cryptocurrencies (i.e., Bitcoin, Bitcoincash, Ethereum, Litecoin, Ripple, and Tether) and the VCRIX using wavelet techniques.

Artificial intelligence techniques

Advances in artificial intelligence technology offer new methods for option pricing and hedging. Carvalho et al. (2019) review machine learning methods applied to predictive maintenance. The main methods they review include K-means, random forest (RF), support vendor machines (SVM), and artificial neural networks (ANNs). ANNs are one of the most commonly applied machine learning algorithms but they require huge datasets to learn correctly and avoid overfitting. Ruf and Wang (2020) provide a comprehensive review of studies on the application of neural networks to option pricing. Long Short-Term Memory (LTSM) has been applied to modeling and predicting the entire implied volatility surface of S&P 500 options (Cheng & Zhang 2019; Bloch & Book 2021). The feedback structure of LSTM allows it to characterize long-memory effects in a financial time series, especially for the volatility of financial assets. The K-means method is a popular clustering algorithm that uses an unsupervised strategy to determine a set of clusters. SVM is a widely known method of performing classification and regression tasks and has high precision in separating different classes of data (Susto el al. 2013). As a supervised learning algorithm, the RF method has been one of the most popular research methods for data mining and is widely used in classification and regressions (Liu, Wang and Zhang, 2012). With RF, the input data pass through multiple decision trees. The algorithm constructs a different number of decision trees at training time and outputs the classification scheme or mean prediction for a regression analysis of the individual data points. RF offers high classification accuracy with tolerable outliers and noise. While deep decision trees may suffer from overfitting, RF avoids overfitting in most cases because it works with random subsets of features and builds smaller trees from those subsets (Biau & Scornet 2016). Hybrid techniques have been employed to build predictive models that combine one or more of the abovementioned techniques. Salamai (2023) proposes a hybrid deep learning approach for efficient predictive modeling of daily and weekly crude oil prices, AdaBoost Random Forest is used for predictive modeling of high-frequency series, and multi-path spatial–temporal deep learning has been used in predictive models of low-frequency series.

Methodology in this study

Dynamic approaches to modeling an IVS cannot fully address the time variation observed in actual IVS. If the coefficients must be regularly reestimated (e.g., every year, week, or day), the model will only perform well if the estimated coefficients remain the same in the next time period, but empirically this is often not the case. Instead of pursuing the dynamic IVS approach, this study is based on the deterministic DIVF approach. DFW suggest that the deterministic volatility framework could be generalized. For example, the volatility surface may be related to past changes in an asset's price. Such a generalized volatility surface is probably the last candidate model that can be considered before resorting to fully stochastic volatility processes, which are difficult to estimate and do not permit option valuation based on a "no arbitrage" condition. Referring to DFW's suggestion that the IVS may be related to past changes in the index level, we propose to include a momentum indicator, which reflects past changes in the underlying asset's price, as well as investor sentiment that is not considered in the BSM model but has a significant impact on option prices. Amin et al. (2004) find that investors' supply of and demand for options may be affected by their return expectations based on market momentum. A momentum indicator, a popular type of trend indicator, is often included in investor sentiment models (Kim & Ryu 2021; Ryu et al. 2017; Chen, Chong & Duan 2010). From a behavioral finance perspective, Barberis et al. (1998) distinguish between two "states" or "regimes" in their model of investor sentiment. In the first state, investors believe that asset returns are mean-reverting, whereas in the second state the returns trend may be likely to rise further after an increase. Investors' points of view alternate between regimes depending on market conditions.

RSI is a popular momentum indicator to measure the speed and magnitude of an asset's recent price changes, and is often used in technical analysis to determine whether an asset is overbought or oversold. It measures the ratio of up-moves to down-moves and normalizes the result so that the value is in a range of 0 to 100. Traditionally, an RSI reading of ≥ 70 indicates an overbought situation. A reading of ≤ 30 indicates an oversold condition. RSI is defined as:

$$\text{RSI} = 100 - \frac{100}{1 + RS} \tag{1}$$

where RS is the relative strength given by:

$$\text{RS} = \frac{\textit{Average gain over the designated period}}{\textit{Average loss over the designated period}} \tag{2}$$

When implementing the RSI, Welles (1978) recommends using a 14-day calculation period, while Petitjean (2004) suggests determining the optimal period via a regression based on the trading style of the investor. In practice, relative strength is often calculated using 14 trading days of price data (RSI-14d), but an RSI can be calculated using various resolutions (time frames). For example, RSI-14d is calculated using 14 trading days of price data, and RSI-14 h is calculated using 14 trading hours of data. For options trading, a longer time frame, measured in days or weeks, may be more relevant for options with longer expiries while a shorter time frame, measured in hours, may be more appropriate for options that are close to expiring. This corresponds well with past studies involving short- and long-term components of option price variance, as discussed in the literature review (Christoffersen et al. 2008; Guo et al. 2014).

To the best of our knowledge, RSI as a momentum indicator has not been included as a predictive factor in modeling IV in prior studies. Though spot prices, spot price volatility and volatility indexes have been suggested as explanatory factors in the dynamic approach, such measures of volatility do not distinguish bullish trends from bearish trends, whereas RSI provides information about the market direction. As RSI is usually available in real time, data can be collected for models, and real-time predictions can be made using these models. RSI can be calculated using different time resolutions as short, medium, and long-term indicators.

To be clear, we are not attempting to forecast option prices or implied volatility over a specific time horizon. Numerous studies have been conducted to predict option price (Venter et al. 2020) and volatility/implied volatility using a time-series model (Fengler 2009; Guo & Han 2018; Hoang & Baur 2020; Chi & Hao 2020). Instead, we provide a generalized deterministic IV function as an alternative to existing option pricing models. The main application of our model is risk management. One of the most widely used methods of portfolio risk management is the Value-at-Risk (VaR), which summarizes the expected maximum loss for a given portfolio over a target horizon at a given confidence level. Within the Basel III regulatory framework for international banks, a VaR risk model may be used to measure the risk of the bank's assets. Junior et al. (2022) employ Generalized Autoregressive Score and Generalized Autoregressive Conditional

Heteroskedasticity (GARCH) models to compute VaR for precious metals. Similarly, our model can facilitate risk management by providing better estimates of option prices in worst-case future scenarios, specifically focused on more accurate VaR calculations for cryptocurrency portfolios. Such estimations are important for margin calculations and preparing collateral for options trading. The RSI indicator used in our model allows risk managers to understand the degree of market stress under various scenarios. For example, the model can answer questions such as, "how will the price of a put option with one month to expiry and a strike price 30% below the current asset value change if the market suddenly drops (e.g., RSI-14d and RSI-14 h are both around 30) and two weeks later the underlying has lost 20% of its value?" The change in the put option's price will be different if the underlying's price does not drop suddenly but gradually declines by 20% over two weeks. With better estimates under such scenarios, risk managers and traders can better model a portfolio's VaR over the option's full life cycle.

A stochastic volatility model using GARCH is comparable to our approach as it estimates option values based on the current spot price and observed path of historical spot prices (Heston and Nandi, 2000) and offers a closed-form solution. However, the advantage of our model is RSI's ability to distinguish between bear and bull markets and its availability in multiple resolutions. Compared to a GARCH model's coefficients, RSI can be easily understood and applied by risk managers in devising stress tests for different types of assets.

Model specification

Our model is mainly based on the DFW's original and subsequently refined DIVF model. Figure 2 is a high-level overview of our theoretical framework. RSI is one of the factors included in our model, where it is used as a bearish indicator (RSI ≤ 35) and a bullish indicator (RSI ≥ 65). We also compare linear regression methods against the RF machine learning algorithm.

Four models and their results are listed below to illustrate the contribution of RSI and the use of the machine learning regression algorithm.

Model 1: DVIF model without RSI, using linear regression
Model 2: DVIF model without RSI, using machine learning RF regression
Model 3: DVIF model with RSI added, using linear regression
Model 4: DVIF model with RSI added, using machine learning RF regression

In discussing the models, the following abbreviations are used:

K: Strike Price, S: Current Price (of the underlying), M: Moneyness (= K/S)*, T: Time to Maturity, N: Proportional Moneyness ($=\text{Log}(M)/\sqrt{T}$), RSI-H: RSI-14 on hourly basis, RSI-H-BEAR: min(0, 0.35 – RSI-H), RSI-H-BULL: min(0, RSI-H – 0.65), RSI-D: RSI-14 on daily basis, RSI-D-BEAR: min(0, 0.35 – RSI-D), RSI-H-BULL: min(0, RSI-D – 0.65).

*Moneyness as defined in Daglish et al. (2007) and Cont & Fonseca (2002).

Model 1 Alentorn model without RSI, using a linear regression (Alentorn 2004):

$$\text{IV}(\text{N},\text{T}) = \beta_0 + \beta_1\text{N} + \beta_2\text{N}_2 + \beta_3\text{T} + \beta_4\text{T}_2 + \beta_5\text{NT} \tag{3}$$

This model is derived from DFW's model by replacing strike price with maturity-adjusted proportional moneyness, as adopted in Rosenberg (2000).

As explained in Alentorn (2004), the volatility smile across moneyness is captured by a second-order function that represents the parabolic shape of the smile. The β_1 coefficient controls for the displacement of the origin of the parabola with respect to the at-the-money options, while the β_2 coefficient controls for the width of the smile. The coefficients β_3 and β_4 control for the effect of time to expiration, i.e., the volatility term structure.

Model 2 same as Model 1, using RF regression.

Model 3 Includes RSI, using linear regression.

The parameter RSI is included as one of the factors and is separated into two components, bearish and bullish. Generalized linear regression is used to obtain the coefficients of each variable. Compared to Model 1, only the second-order terms of proportional moneyness and time to maturity are retained, to reduce collinearity.

$$\begin{aligned}
&\mathrm{IV(N, T, RSI-h-bear,\ RSI-h-bull, RSI-d-bear,\ RSI-d-bull)} = \beta 0 + \beta_1 \mathrm{N}_2 + \beta_2 \mathrm{T}_2 + \beta 3\mathrm{NT} \\
&\quad + \mathrm{B}_4 \mathrm{RSI-d-bear} + \beta_5 \mathrm{RSI-d-bull} + \beta_6 \mathrm{RSI-d-bear} * \mathrm{N} + \beta_7 \mathrm{RSI-d-bull} * \mathrm{N} \\
&\quad + \beta_8 \mathrm{RSI-h-bear} + \beta_9 \mathrm{RSI-h-bull} + \beta_{10} \mathrm{RSI-h-bear} * \mathrm{N} + \beta_{11} \mathrm{RSI-h-bull} * \mathrm{N}
\end{aligned} \tag{4}$$

Model 4 same as Model 3, using RF regression.

Therefore in Model 4, all of the factors in Model 3 (N,T, RSI-h-bear, RSI-h-bull, RSI-d-bear, RSI-h-bull) are included and RF is used as a non-linear regression.

The statistical analysis was conducted using the SAS software package. Invoking the RF algorithm and plotting the implied volatility 3-D surface were coded in Python.

Data collection

In this study, we model IV for Bitcoin. Cryptocurrencies often exhibit high volatilities and therefore provide richly diverse data and significant challenges for model development and testing. In this study, we use options data from the largest cryptocurrency options exchange Deribit® (www.deribit.com). We downloaded Deribit option chain data for the period from June 3, 2022 to July 31, 2022 as the in-sample (training) dataset, and data for the period from August 1, 2022 to September 30, 2022 as the out-sample (validation) dataset. The data were obtained at 0:00 h and 12:00 h UST every day. The option chain data contains the implied volatility calculated for each option contract (assuming a 0% interest rate). Real-time data is freely available from the Deribit website and therefore we began to collect daily data immediately after this research study was conceptualized. We chose two months of data collection for training and another two months for validation as we believe these are the minimal time periods necessary to obtain meaningful results from this exploratory study. Historical option chain data over longer time horizons are available by subscription but the authors do not have access to such data; hence development/validation of the model over longer time horizons awaits future study.

Similarly, for RSI we access data from the website www.taapi.io to obtain RSI values at 0:00 h and 12:00 h UST every day. This real-time data is also freely available. The standard number of periods used to calculate the initial RSI value is 14. We obtained RSI values daily for RSI-14d (momentum over 14 days, to reflect medium-term market

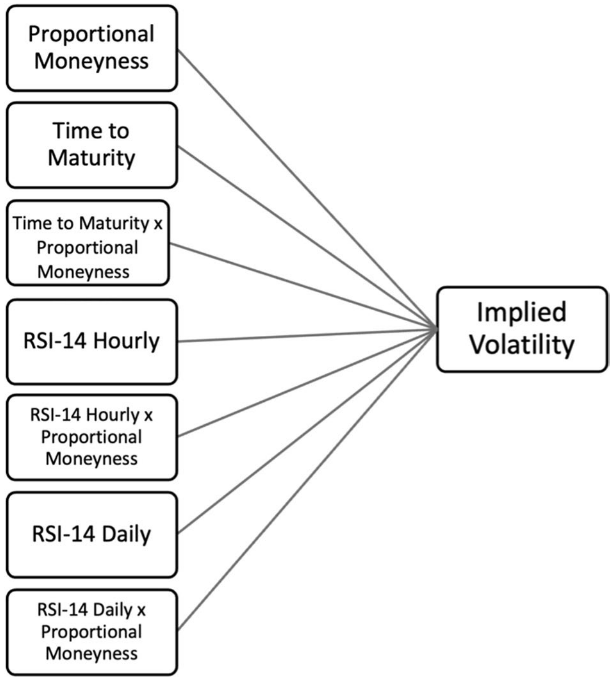

Fig. 2 High level overview of theoretical framework in this study

momentum) and hourly for RSI-14 h (momentum over 14 h, to reflect the short term market momentum). For comparison, past studies on securities and currencies were often constrained by data availability due to illiquidity. For example, the IVS of options on securities and currencies could only be constructed from a small number of maturities, such as 1, 2, 3, 6, 9, 12, 18, and 24 months to expiry from the date of issue. In comparison, there are maturities of 2 days, 1 week, and 1, 2, 3, 6, 9, and 12 months for Bitcoin option.

We collected data only for options for which both bid and ask prices were available. IV was obtained using the mid-point of the bid-ask quote. Some prior studies only include options that fall within a certain range of moneyness. For example, Heston and Nandi (2000) only consider options where the absolute difference between the strike price and the underlying price is $\leq 10\%$, while Cont and Fonseca (2002) use a cut-off of $\leq 50\%$, as the numerical uncertainty of IV outside this range could be considered too high. Most studies remove in-the-money-options and use only out-of-the-money options (Wang et al. 2017) as out-of-money options are more sensitive to volatility and are more actively traded. In the current study, we also exclude in-the-money options but include all out-of-the-money options regardless of moneyness. We consider it too restrictive to exclude far out-of-money options for cryptocurrencies.

In terms of time to expiry, prior studies exclude options with expiries of less than a few days (6 days in Heston & Nandi (2000) and DFW (1998); seven days in Wang et al. (2017)) as those very short options are less liquid and a minor bias in their prices could lead to a substantial distortion in IV (Wang et al 2017). However, considering that short maturity options for cryptocurrencies are sometimes heavily traded, options with less than seven days to expiry were included in our study; only those with less than one day to expiry were excluded. Long maturity options, i.e., those expiring in more than 90 days, were excluded as their low liquidity might distort prices. Other studies also exclude long maturities (e.g., Heston & Nandi (2000), DFW (1998) exclude options with a time to expiry greater than 100 days).

To transform this data, the time to expiry in days was divided by the maximum time to expiry (in our case, 90 days). Values of RSI, which were originally scaled from 1 to 100, were further divided by 100. Regarding parameter setting for the RF algorithm, we set the maximum depth of a tree equal to eight and the number of estimators to 15 to avoid overfitting.

Results

Descriptive statistics

We collected data for four months, resulting in a reasonable number of data points due to the higher liquidity of cryptocurrency options and their 24-h trading cycle. Table 2 provides descriptive statistics of the data collected for Bitcoin put and call options for the in-sample (training) and out-sample (validation) datasets, from which we observe the lowest strike of the bid put options is merely 20% of its current price and the highest strike price of the bid call options is 585% of its current price. This implies that some traders are betting that Bitcoin will fall dramatically while others are betting that Bitcoin will rise dramatically within the next 90 days. This illustrates the extreme volatility of Bitcoin expected by traders. Compared to call options, put options have a wider range of IV, likely because Bitcoin was in a bear market during the period studied and traders emphasized the possibility of a significant price drop in these options more heavily over this period. For example, the standard deviation of IV for put options in the training dataset is 32%, compared to 15% for call options.

Figure 3 plots the IVS. We note that IV covers a wider range for near-expiry options (time of expiry less than six days) and depends on the moneyness of the option. When options are close to expiry, their prices may be much higher during a volatile market (hence higher IV) versus a quiet market. In a volatile market, the price of Bitcoin may suddenly jump up or down, moving closer to the option's strike price; a risk premium is therefore reflected in the option price quote.

Compared to previous studies of IV for securities and currencies, IVs for cryptocurrencies are generally much higher. For example, Guo et al. (2014) report that the mean of 30-day volatility of S&P 500 index call options from January 2005 to October 2010 was 19.64% while Bitcoin call options in our study period had a mean 30-day volatility of 81%. Therefore, one should exercise caution in comparing regression model results and errors across different asset types as larger errors are expected for assets with higher volatilities, even though the models may have the same explanatory power.

Regression results

Here we re-list our four models, and their regression results are then discussed.

Model 1: DVIF model without RSI, using linear regression
Model 2: DVIF model without RSI, using machine learning RF regression
Model 3: DVIF model with RSI added, using linear regression
Model 4: DVIF model with RSI added, using machine learning RF regression

To investigate the contribution of RSI, we compare performance between comparable models with and without the RSI indicator (Model 1 versus Model 3; Model 2 versus Model 4). On the other hand, to investigate the contribution of the nonlinear machine learning algorithm RF, we compare the results of the comparable models that use linear regression versus RF (Model 1 versus Model 2; Model 3 versus Model 4). The adjusted R^2 and root mean square error (RMSE) are shown where appropriate (see Table 3). Table 4 shows Model 3's regression coefficients and Fig. 4 plots the predicted IV against the actual IV.

In general, models that include RSI perform better than models without RSI (i.e., Model 3 performs better than Model 1, and Model 4 performs better than Model 2), especially for put options. For put options, the results of Model 3, which uses a linear regression, produce an adjusted R^2 of 0.79 for the in-sample dataset, compared to an R^2 of 0.60 for Model 1. This indicates that RSI can help to explain the IV variance in bullish versus bearish markets. With respect to the RF algorithm, by comparing results between model 4 and model 3 we see that RF performs much better than linear regression for the in-sample dataset, reducing RMSE by 50%. For example, using the in-sample dataset of put options, RMSE for Model 3 is 0.1481 but is only 0. 0601 in Model 4. However, the RF algorithm does not consistently outperform linear regression with the out-of-sample dataset. This may be due to common overfitting problems in machine learning algorithms; such problems may be overcome by using a more extended period of data for training and validation. Overall, our models, which incorporate RSI and use the machine learning RF algorithm, significantly outperform the DIVF model derived from DFW, reducing the RMSE by approximately 20%–40%.

Table 2 Descriptive statistics showing data distribution in the training and validation for Bitcoin put and call option for data period Jun 2022–Sep 2022

Variable	Bitcoin put							
	Training (N = 7609)				Validation (N = 6651)			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Impliedvolatility(%)	101	32	46	466	83	17	39	161
Moneyness	0.80	0.14	0.20	0.99	0.80	0.16	0.21	0.99
Timetoexpiry(day)	24.84	23.94	1	90	28.74	26.05	1	90
RSI-d	0.40	0.11	0.19	0.63	0.46	0.09	0.29	0.62
RSI-h	0.49	0.12	0.19	0.84	0.48	0.12	0.11	0.67
Variable	**Bitcoin Call**							
	Training(N = 10,032)				**Validation(N = 8061)**			
	Mean	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
Impliedvolatility(%)	81	15	45	275	71	9	36	129
Moneyness	1.46	0.56	1	5.85	1.33	0.34	1	3.94
Timetoexpiry(day)	32.39	26.02	1	90	33.3	26.2	1	90
RSI-d	0.39	0.11	0.19	0.63	0.45	0.09	0.29	0.62
RSI-h	0.48	0.12	0.19	0.84	0.47	0.12	0.11	0.78

N number of data, *SD* standard deviation, *Min* minimum, *Max* maximum

Robustness test

We conducted a robustness check for Model 3, the linear regression model that includes RSI. First, we use a variance inflation factor (VIF) analysis to check for collinearity. As a rule of thumb, a variable whose VIF values are greater than 10 may require further investigation. As shown in Table 3, all of the estimators have a VIF < 10; thus, collinearity is not a concern. For heteroscedasticity, we examine the plot of the residuals to determine whether they have a constant variance (Fig. 5). When examining the residuals versus predicted values and Q-Q (quantile-to-quantile) residuals plot, we observe heteroscedasticity as the variance is asymmetrical with more positive outliners. That is, the model underestimates IV in some cases. The reason may be that in periods of market stress, option prices (and hence the IV) may suddenly jump higher. This hypothesis is supported by plotting the residuals on a graph whose axes are time-to-expiry and proportional moneyness. The residuals show more positive outliers for "near the money" options (options with moneyness approaching 1) and when the time to expiry is < 0.1 (that is $90 \times 0.1 = 9$ days, as our maximum time to expiry is 90 days). This may be why most previous studies exclude option contracts that expire in less than six days; if such short maturity options are excluded, the residual plot would be close to normal. This highlights the importance of exploring the use of nonlinear machine learning algorithm (in Model 4) as such nonlinearity may be better addressed using a nonlinear model.

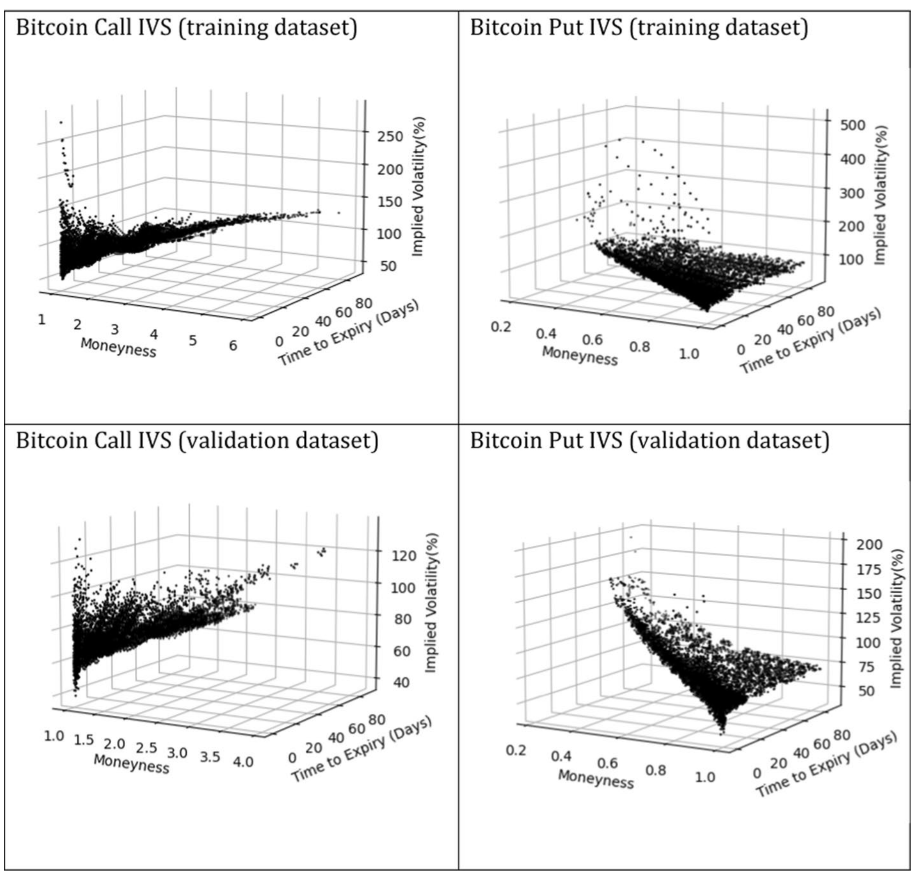

Fig. 3 Implied volatility surfaces (IVS) for Bitcoin put and call options

One challenge, as mentioned earlier, is that the study's results may only be applicable to the specific four-month period studied. Although there was a significant price fluctuation during this period (with Bitcoin prices dropping from USD 29,000 to around USD 18,000, then rebounding to USD 24,000), additional evidence may be required to validate our model for a longer duration. Although accessing historical data typically requires a subscription, the data provider allowed us to download 24-h of Bitcoin option data for the first day of each month since April 2019. We extracted the first and last observations for the first day of the month from April 2019 to March 2022 and reexamine our model based on this data. Volatility over this extended period was higher than in our original period and showed a larger range of RSI reflecting bear and bull market runs (Table 5). For example, for this extended data period the maximum RSI is 92 while the maximum RSI in our original dataset is 68. Despite this, using linear regression analyses based on two models (with and without RSI), we show that Model 3, which includes RSI (Table 6) maintains an advantage over Model 1, which does not include RSI. For put options, the R^2 for Model 1 is 0.70 while for Model 3 the R^2 is 0.76. For call options, the R^2 for Model 1 is 0.50 while for Model 3 it is 0.56. These results suggest that a momentum indicator is a significant factor in accounting for option price variance over a longer time period.

Discussions and conclusion

The DIVF developed in this study presents an easy-to-compute estimation of IV for any strike price or time to expiry, provided the underlying momentum is known. Based on our results, this study makes the following contributions:

Table 3 Model 3 Estimated Parameters and VIF (Model 3 – Training dataset) for data period Jun 2022– Sep 2022

Parameter estimates

Label	DF	Parameter estimate	Standard error	*t* value	Pr > |t|	Tolerance	Variation inflation
BTC call							
Intercept	1	0.68549	0.00132	521.07	<.0001		0
N (Proportional moneyness)*N	1	0.26168	0.00243	107.56	<.0001	0.36692	2.72537
N*t	1	− 0.15179	0.00433	− 35.05	<.0001	0.49607	2.01585
RSI-14d bearish	1	2.22801	0.03503	63.61	<.0001	0.29392	3.40225
RSI-14d bearish*N	1	− 1.41921	0.04751	− 29.87	<.0001	0.24072	4.15417
RSI-14 h bullish	1	− 0.09112	0.3660	− 2.49	0.0128	0.98347	1.01680
BTC put							
Intercept	1	0.81471	0.00272	299.81	<.0001		0
t*t	1	− 0.22098	0.00890	− 24.83	<.0001	0.59347	1.68501
N (Proportional moneyness)*N	1	0.18088	0.00289	62.65	<.0001	0.39196	2.55127
N*t	1	− 0.35026	0.01385	− 25.28	<.0001	0.55776	1.79288
RSI-14d bearish	1	2.51965	0.06667	37.79	<.0001	0.37254	2.68428
RSI-14d bearish*N	1	− 2.25133	0.08903	− 25.29	<.0001	0.25499	3.92168
RSI-14 h bullish	1	− 0.69910	0.11488	− 5.82	<.0001	0.34816	2.87249
RSI-14 h bearish	1	0.25272	0.05304	4.76	<.0001	0.87781	1.13920
RSI-14 h bullish*N	1	− 0.94245	0.16274	− 5.79	<.0001	0.34305	2.91506

BTC: Bitcoin; K: Strike price; S: Current price; M: Moneyness (= K/S); T: Time to maturity; N: log(M)/$\sqrt{T}$

Table 4 Tabulation of statistical variation due to regression for different model. Model 4 has the least value of RMSE both in the put and call options among the four models, indicating that the Model 4 is the most optimal method to model the IV with underlying RSI and with RF as the non-linear regression method

	Bitcoin				
	Put option				
	Adjusted R^2	**RMSE**	**Average IV**	**RMSE**	**Average IV**
Model	(in-sample)	(in-sample)	(in-sample)	(out-sample)	(out-sample)
1	0.6071	0.2060	101%	0.1510	84%
2	N.A	0.1640	101%	0.1509	84%
3	0.8117	0.1426	101%	0.0963	84%
4	N.A	0.0601	101%	0.0930	84%
Bitcoin					
	Calloption				
Model	**AdjustedR^2**	**RMSE**	**AverageIV**	**RMSE**	**AverageIV**
1	(in-sample)	(in-sample)	(in-sample)	(out-sample)	(out-sample)
2	0.5409	0.1025	82%	0.0913	71%
3	N.A	0.0853	82%	0.0948	71%
4	0.6933	0.0843	82%	0.0722	71%
	N.A	0.0416	82%	0.0667	71%

Model 1: linear regression without RSI, Model 2: RF regression without RSI, Model 3: linear regression with RSI, Model 4: RF regression with RSI

RMSE: Root mean square error; IV: implied volatility; RF: Random Forest machine learning regression method

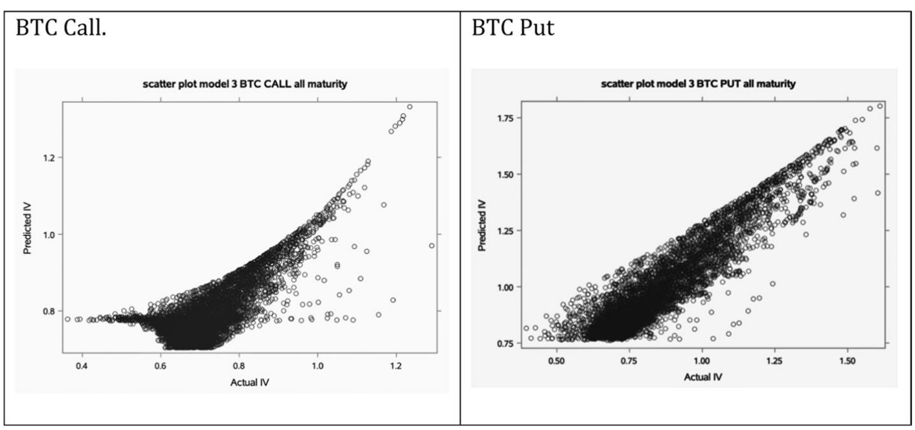

Fig. 4 Plot of predicted IV vs actual IV (Model 3 – Training dataset). BTC: Bitcoin, IV: implied volatility

1. Including the RSI momentum indicator as a factor enhances the accuracy of the deterministic function for modeling IV. Our model captures both short-term and medium-term momentum through RSI momentum indicators calculated over two different timeframes. It shows that the deterministic implied function approach, if enhanced by certain factors, may still be relevant for option pricing due to its relatively simple formulation and ease of computation compared to prevailing stochastic option pricing models. As noted in DFW, a stable DIVF option valuation model offers an important way to compute hedge ratios and value exotic options.

2. As a real-time momentum indicator, the RSI is readily available from trading websites and software, so our DIVF model can be easily deployed for real-time calculations.
3. The model exhibits stable in-sample and out-sample performance. The parameters, estimated during model development, remain the same in estimating option prices over the next two months. We expect that our model will not require frequent calculation iterations to obtain convergence for the coefficients, while stochastic option pricing models and dynamic IVS models require frequent reestimation of parameters. In addition, inputting option characteristics or initial conditions, such as at-the-money known IV or IVS as in the dynamic approach, is not required.
4. By using a nonlinear machine learning regression algorithm to replace the linear regression algorithm used in prior DIVF studies we achieve better IVS modeling and IV predictions. Nonlinear modeling may be particularly relevant for "near-the-

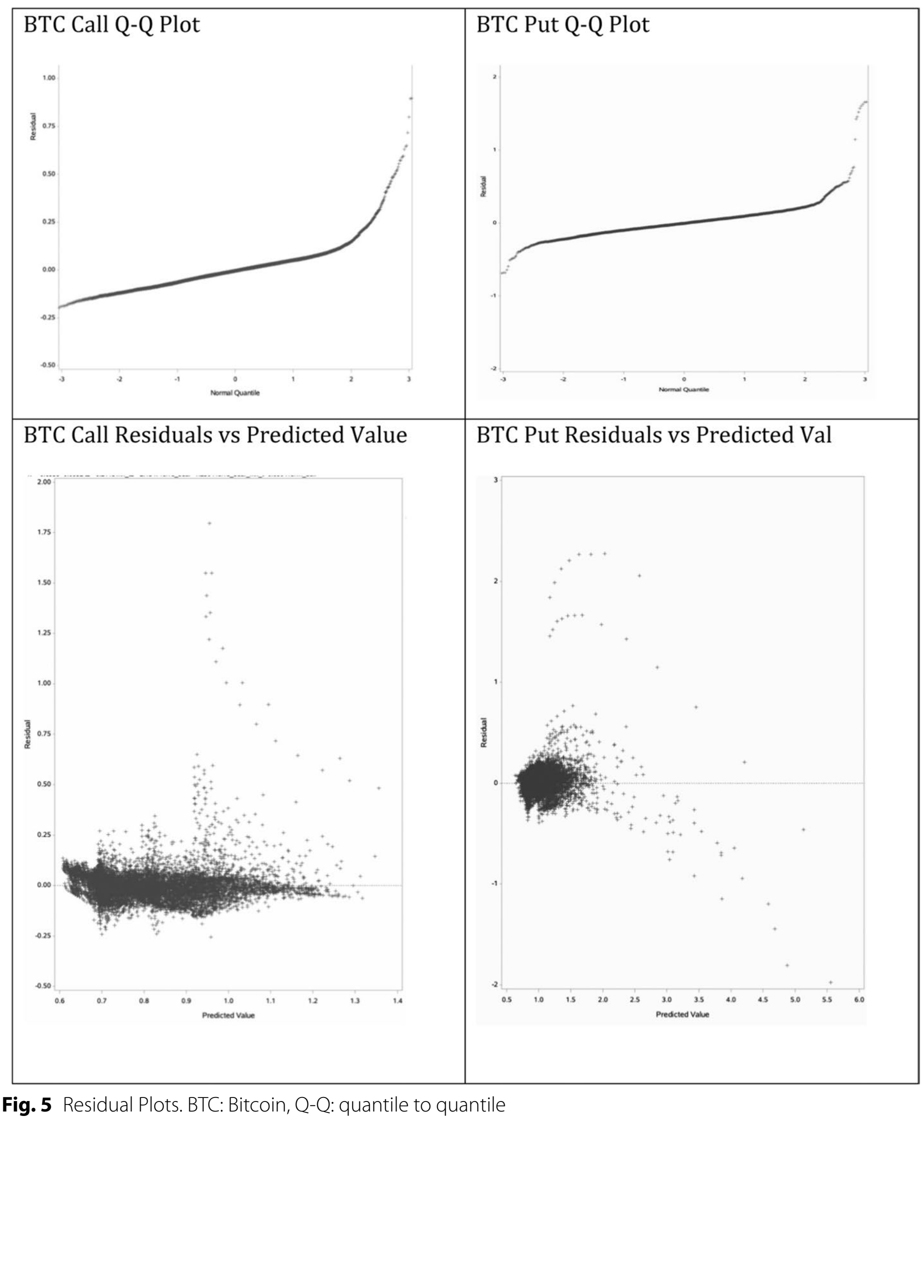

Fig. 5 Residual Plots. BTC: Bitcoin, Q-Q: quantile to quantile

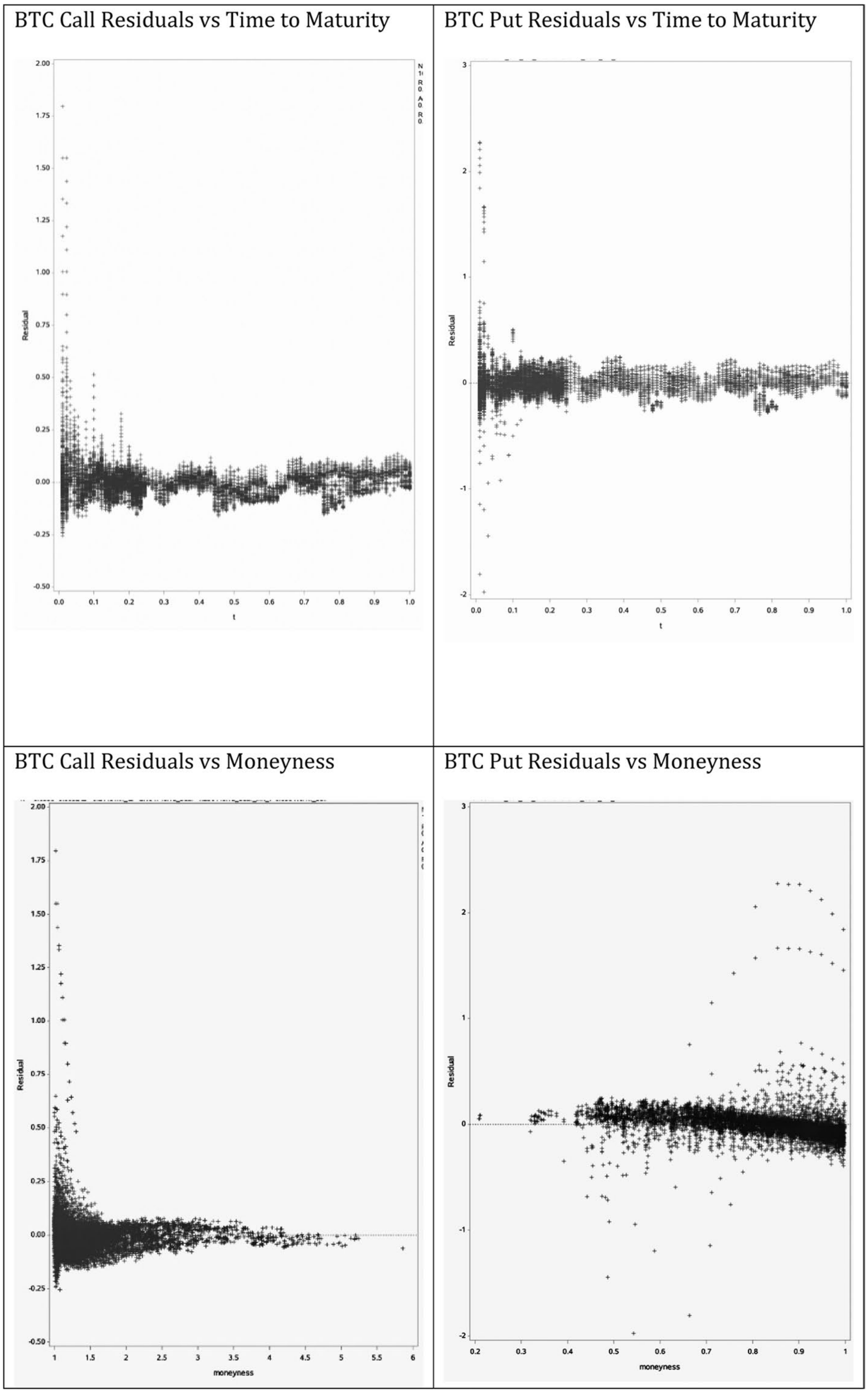

Fig. 5 continued

money" options or options with short maturities, as sudden price jumps for these options during market stress cannot be captured by a linear model.

5. The study uses data on Bitcoin options, which is fairly uncommon. The highly volatile nature of the Bitcoin market is challenging to model-building, and our results indicate that our model performs well even using such volatile market data.
6. Unlike prior studies, very short-term maturities and a full range of moneyness are included in our option chain data and therefore, our model can be relevant to trading options that are close to expiration.
7. Although our study does not intend to predict the evolution or model the time dynamics of an IVS, the price path of an underlying asset can be forecasted using different models through Monte Carlo or other means, and, by calculating RSI, the entire IVS can be estimated by our model.

However, there are limitations to our study, and there is room for further optimization.

1. Future studies can compare the performance of our model and other option pricing models, such as DFW's ad-hoc BSM model of Heston & Nandi (2000).
2. This study was conducted using Bitcoin, a cryptocurrency with a different IVS compared to traditional assets such as listed securities or currencies. Empirical tests of our model using options on traditional assets would be interesting.
3. Although the in-sample performance of the nonlinear machine learning RF algorithm is excellent, its out-sample performance is only marginally better than the results from using a linear regression. This may be due to the overfitting issues usually encountered with machine learning algorithms which may be overcome by using a more extended data collection period, which future studies could explore.
4. During our in-sample data period, cryptocurrencies were mainly in a bear market and the RSI on daily basis did not exceed 70. Yet the RSI on an hourly basis includes both bullish and bearish values, and such data diversity should satisfy the requirements for model development and testing. In our extended data period covering three years, which included bear and bull markets, the RSI computed on daily basis exceeded 90 and the momentum indicators remained significant. Therefore, our methodology should be sufficiently applicable to markets under both bull and bear

Table 5 Descriptive statistics showing data distribution for Bitcoin put and call option for the extended data period Apr 2019 – Mar 2022

	Bitcoin put				**Bitcoin call**			
	N = 2168				**N = 2565**			
Variable	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
Implied voltality (%)	79	15	40	158	78	15	34	293
Moneyness	0.84	0.22	0.17	1.67	1.29	0.49	0.36	4.78
Time to expiry (day)	29.11	22.24	1	88	28.25	21.79	1	88
RSI-d	0.55	0.18	0.18	0.92	0.55	0.18	0.18	0.92
RSI-h	0.54	0.14	0.26	0.86	0.54	0.14	0.26	0.86

N number of data, *SD* standard deviation, *Min* minimum, *Max* maximum

Table 6 Model 3 Estimated Parameters for extended data period Apr 2019 – Mar 2022

Parameter estimates

Label	DF	Parameter estimate	Standard error	*t* value	Pr > |t|	Tolerance	Variation inflation
BTC call							
Intercept	1	0.67202	0.00314	214.3	< 0.001		0
t*t	1	0.08013	0.01233	6.50	< 0.001	0.65096	1.53619
N*N	1	22.14118	0.65794	33.65	< 0.001	0.36235	2.75980
N(Proportional moneyness)*t	1	− 0.00329	0.00167	− 1.96	0.0499	0.27432	3.64533
RSI-14 h bullish	1	0.39090	0.05545	7.05	< 0.001	0.58488	1.70975
RSI-14 h bearish	1	2.25287	0.16208	13.90	< 0.001	0.61360	1.62972
RSI-14d bearish	1	0.62513	0.07742	8.07	< 0.001	0.62814	1.59199
RSI-14 h bearish*N	1	− 8.65585	2.45402	− 3.53	0.0004	0.59897	1.66952
RSI-14 h bullish*N	1	− 1.66548	0.86697	− 1.92	0.0548	0.15422	1.94392
RSI-14d bearish*N	1	− 6.14497	1.18733	− 5.18	< 0.001	0.59778	1.67286
BTC put							
Intercept	1	0.66313	0.00257	257.89	< .0001		0
t*t	1	0.09649	0.00960	10.05	< .0001	0.67545	1.48049
N*N	1	26.17348	0.52531	49.82	< .0001	0.36112	2.76915
N (Proportional moneyness)*t	1	0.0054	0.00134	4.15	< .0001	0.28723	3.48149
RSI-14 h bullish	1	0.31272	0.03430	9.12	< .0001	0.91424	1.09381
RSI-14 h bearish	1	1.93549	0.12765	15.16	< .0001	0.63095	1.58491
RSI-14d bearish	1	0.76817	0.05111	15.03	< .0001	0.91585	1.09189
RSI-14 h bearish*N	1	12.32127	2.01088	6.13	< .0001	0.60621	1.64960
RSI-14d bearish*N	1	− 0.76296	0.34078	− 2.24	0.0253	0.86036	1.16231

BTC: Bitcoin; K: Strike price; S: Current price; M: Moneyness (= K/S); T: Time to maturity; N: log(M)/$\sqrt{T}$

conditions. A longer data collection period with model development and validation using a moving window may further verify the contribution of momentum indicators in our model.

Abbreviations

ANN	Artificial neural networks
BSM	Black–Scholes-Merton
BTC	Bitcoin
DFW	Dumas B., Fleming J., Whaley R.
DIVFs	Deterministic implied volatility functions
GARCH	Generalized autoregressive conditional heteroscedasticity
IV	Implied volatility
IVS	Implied volatility surface
K	Strike price
Q-Q	Quantile-to-quantile
RF	Random forest
RS	Relative strength
RSI	Relative strength index
RSI	Relative strength index
S	Asset price
SVM	Support vendor machines
VaR	Value-at-Risk
VIF	Variance inflation factor
VIX	Volatility Index

Acknowledgements
None.

Author contributions
FL: data collection, programming and analysis, ML: programming and analysis and independent checking, SKD: project proposal, overall project management and result countercheck. All authors agreed on the final results.

Funding
None.

Availability of data and materials
From website download and self developed technique.

Declarations

Competing interests
None.

Received: 10 October 2022 Accepted: 7 February 2024
Published: 31 December 2024

References

Agyei SK, Adam AM, Bossman A, Asiamah O, Junior PO, Asafo-Adjei R, Asafo-Adjei E (2022) Does volatility in cryptocurrencies drive the interconnectedness between the cryptocurrencies market? Insights from wavelets. Cogent Econ Financ 10(1):2061682
Alentorn A (2004) Modelling the implied volatility surface: an empirical study for FTSE options. University of Essex
Amin K, Coval JD, Seyhun HN (2004) Index option prices and stock market momentum. J Bus 77(4):835–874
Arunsingkarat S, Costa R, Misran M, Phewchean N (2021) Option pricing under GARCH models applied to the SET50 index of Thailand. WSEAS Trans Math 20:112–121
Asafo-Adjei E, Junior PO, Adam AM (2021) Information flow between global equities and cryptocurrencies: a VMD-based entropy evaluating shocks from COVID-19 pandemic. Complexity 2021:1–25
Bakshi GS, Cao C, Chen ZW (1997) Empirical performance of alternative option pricing models. J Financ 52:2003–2049
Balcilar M, Bouri E, Gupta R, Roubaud D (2017) Can volume predict Bitcoin returns and volatility? A quantiles-based approach. Econ Model 64:74–81
Barberis N, Schleifer A, Vishny R (1998) A model of investor sentiment. J Financ Econ 49(1998):307–343
Barson Z, Junior PO, Adam AM, Asafo-Adjei E (2022) Connectedness between gold and cryptocurrencies in COVID-19 pandemic: a frequency-dependent asymmetric and causality analysis. Complexity 2022:1–17
Bates DS (2000) Post-'87 crash fears in the S&P 500 futures option market. J Econom 94(1–2):181–238
Biau G, Scornet E (2016) A random forest guided tour. TEST 25:197–227
Black F, Scholes M (1973) The pricing of options and corporate liabilities. J Polit Econ 81(3):637–654
Bloch DA, Book A (2021) Deep Learning Based Dynamic Implied Volatility Surface, Quant Finance Working Paper
Bouri E, Gil-Alana L, Gupta R, Roubaud D (2018) Modelling long memory volatility in the Bitcoin market: Evidence of persistence and structural breaks. Int J Financ Econ 24(1):412–426
Brooks C, Oozeer MC (2002) Modelling the implied volatility of options on long gilt futures, February 2002. J Bus Financ Account 29(1–2):111–137
Câmara A, Heston SL (2008) Closed-form option pricing formulas with extreme events. J Futur Mark 28:213–230
Carr P, Madan D (1999) Option valuation using the fast Fourier transform. J Comput Financ 2(4):61–73
Carr P, Wu L (2004) Time-changed Levy processes and option pricing. J Financ Econ 71(1):113–114
Carvalho TP, Soares FA, Vita R, Francisco RP (2019) A systematic literature review of machine learning methods applied to predictive maintenance. Comput Ind Eng 137:106024
Chen H, Chong TTL, Duan X (2010) A principal-component approach to measuring investor sentiment. Quantit Financ 10(4):339–347
Chen D, Guo B, Zhou G (2022) Firm fundamentals and the cross-section of implied volatility shape. J Financ Mark 63:10071
Cheng S, Zhang Z (2019) Forecasting Implied Volatility Smile Surface via Deep Learning and Attention Mechanism, Guanghau School of Management, Peking University Working Paper
Chew D, Stewart B (2022) Stewart Myers and the MIT school of real options and capital structure. J Appl Corp Financ 34(3):8–21
Chi Y, Hao W (2020) A Horserace of Volatility Models for Cryptocurrency: Evidence from Bitcoin Spot and Option Markets, Auckland University Graduate School of Management
Christoffersen P, Jacobs K, Ornthanalai C, Wang Y (2008) Option valuation with long-run and short-run volatility components. J Financ Econ 90:272–297
Christoffersen P, Heston S, Jacobs K (2009) The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well? Manage Sci 55:1914–1932
Cont R, Fonseca JD (2002) Dynamics of implied volatility surfaces. Quant Financ 2:45–60
Cox JC, Ross SA (1976) The valuation of options for alternative stochastic processes. J Financ Econ 3(1–2):145–166

Daglish T, Hull J, Suo W (2007) Volatility surfaces: theory, rules of thumb, and empirical evidence. Quant Financ 7(5):507–524
Dumas B, Fleming J, Whaley R (1998) Implied volatility functions: empirical test. J Financ 53(6):2059–2106
Elliot RJ, Chan L, Siu TK (2007) Pricing volatility swaps under Heston's stochastic volatility model with regime change. Appl Math Financ 14(1):41–62
Escobar-Anel M, Rastegari J, Stentoft L (2021) Option pricing with conditional GARCH models. Eur J Oper Res 289(1):350–363
Fama E (1965) The behavior of stock market prices. J Bus 38:34–105
Feng C, Tan J, Jiang Z, Chen S (2020) A generalized European option pricing model with risk management. Physica A 545:123797
Fengler MR (2009) Arbitrage-free smoothing of the implied volatility surface. Quant Financ 9(4):417–428
Fengler M, Hardle W, Mammen E (2005) A dynamic semiparametric factor model for implied volatility string dynamics, SFB 649 Discussion Paper
Forlicz M (2011) Comparison of the behaviour of market option prices in relation to option prices resulting from the black-scholes model during periods of a bull and bear market. Mathematical Economics 7(14):71–81
Goncalves S, Guidolin M (2006) Predictable dynamics in the S&P index options implied volatility surface. J Bus 79(3):1591–1635
Goutte S, Ismail A, Pham H (2017) Regime-switching stochastic volatility model: estimation and calibration to VIX options. Appl Math Financ 24(1):38–75
Guo D (2000) Dynamic volatility trading strategies in the currency option market. Rev Deriv Res 4(2):133–154
Guo B, Han Q, Zhao B (2014) The Nelson–Siegel model of the term structure of option implied volatility and volatility components. J Future Mark 34(8):788–806
Guo B, Han Q, Lin H (2018) Are there gains from using information over the surface of implied volatilities? J Futur Mark 38(6):645–672
Harb E, Bassil C, Kassamany T (2022) Volatility interdependence between cryptocurrencies, equity, and bond markets. Comput Econ 2022:1–31
Harvey CR, Whaley RE (1992) Market volatility prediction and the efficiency of the S & P 100 index option market. J Financ Econ 31(1):43–73
Heston SL (1993) A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev Financ Stud 6(2):327–343
Heston SL, Nandi S (2000) A closed form GARCH option valuation model. The Review of Financial Studies 13(3):585–625
Hoang LT, Baur DG (2020) Forecasting bitcoin volatility: Evidence from the options market. J Futur Mark 40(10):1584–1602
Jalal RNUD, Alon I, Paltrinieri, A. (2021). A bibliometric review of cryptocurrencies as a financial asset. Technology Analysis & Strategic Management, 1–16.
Jeckwerth J, Rubinstein M (1995) Recovering probability distributions from contemporaneous security prices, Working paper, University of California, Berkeley
Junior PO, Tiwari AK, Tweneboah G, Asafo-Adjei E (2022) GAS and GARCH based value-at-risk modeling of precious metals. Resour Policy 75:102456
Kim S (2009) Option pricing with extreme events: using Câmara and Heston (2008)'s Model. Asia Pac J Financ Stud 2:187–209
Kim K, Ryu D (2021) Does sentiment determine investor trading behaviour. Appl Econ Lett 28(10):811–816
Liu Y, Wang Y, Zhang J (2012) New machine learning algorithm: Random Forest. In: International Conference on Information Computing and Applications (ICICA 2012), Information Computing and Applications, pp 246–252
Madan DB, Seneta E (1990) The variance gamma (VG) model for share market returns. J Bus 63(4):511–524
Madan DB, Carr PP, Chang EC (1998) The variance gamma process and option pricing. Rev Financ 2(1):79–105
Mandelbrot B (1966) Forecasts of future prices, unbiased markets, and "martingale" models. J Bus 39(1):242–255
Mehrdoust F, Samimi O (2020) Pricing multi-asset American option with stochastic correlation coefficient under variance gamma asset price dynamic. Ann Financ Econ 15(04):2050015
Mehrdoust F, Noorani I, Hamdi A (2021) Calibration of the double Heston model and an analytical formula in pricing American put option. J Comput Appl Math 392:113422
Merton RC (1976) Option pricing when underlying stock returns are discontinuous. J Financ Econ 3(1–2):125–144
Nakamoto S. (2008). Bitcoin: A peer-to-peer electronic cash system, https: //bitcoin.org/ bitcoin.pdf.
Ncube M (1996) Modelling implied volatility with OLS and panel data models. J Bank Finance 20:71–84
Neumann M, Skiadopoulos G (2012) Predictable dynamics in higher order risk-neutral moments: evidence from the S&P 500 options. J Financ Quant Anal 48:947–977
Nzokem, A. H. (2021). Gamma variance model: Fractional fourier transform (FRFT). In: Journal of Physics: Conference Series (Vol 2090, No 1, p. 012094). IOP Publishing.
Panigirtzoglous N, Skiadopoulos G (2004) A new approach to modeling the dynamics of implied distributions: Theory and evidence from the S&P 500 options. J Bank Finance 28:1499–1520
Pena I, Rubio G, Serna G (1999) Why do we smile? On the determinants of the implied volatility function. J Bank Financ 23(1999):1151–1179
Pena I, Rubio G, Serna G (2001) Smiles, bid-ask spreads and option pricing. Eur Financ Manag 7(3):351–374
Petitjean M (2004) Le Guide du Trader: Méthodes et Techniques de Spéculation Boursière, 1st edn. Dunod
Rosenberg JV (2000) Implied volatility functions: a reprise. J Deriv 7(3):51–64
Rubinstein M (1985) Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active CBOE option classes from August 23, 1976 through August 31, 1978. J Financ 40(2):455–480
Rubinstein M (1994) Implied binomial trees. J Financ 49:771–818

Ruf J, Wang W (2020) Neural networks for option pricing and hedging: a literature review. J Comput Financ 24(1):2020
Ryu D, Kim H, Yang H (2017) Investor sentiment, trading behaviour and stock returns. Appl Econ Lett 24(12):826–830
Salamai AA (2023) Deep learning framework for predictive modelling of crude oil price for sustainable management in oil markets, Expert Systems with Applications, Vol. 211, January 2023
Siu TK, Elliott RJ (2021) Bitcoin option pricing with a SETAR-GARCH model. Eur J Financ 27(6):564–595
Susto GA, Schirru A, Pampuri S, Pagano D, McLoone S, Beghi A (2013) A predictive maintenance system for integral type faults based on support vector machines: An application to ion implantation. In: 2013 IEEE International Conference on Automation Science and Engineering (CASE), Madison, WI, USA, 2013, pp 195–200
Venter PJ, Mare E, Pindza E (2020) Price discovery in the cryptocurrency option market: a univariate GARCH approach. Econ Financ 8:1803524
Voukelatos N, Verousis T (2019) Option-implied information and stock herding. Int J Financ Econ 24:1429–1442
Wang J, Chen S, Tao Q, Zhang T (2017) Modelling the implied volatility surface based on Shanghai 50ETF options. Econ Modell 64:295–301
Welles WJ (1978). New concepts in technical Trading System, Trend Research

Publisher's Note