Algorithm for Adjusting Implied Volatility Curves in the Options Market

September, 2025

Author 1: Thomas Barton
He is a Computer Science undergraduate at Inteli and Data Engineer intern at BCG X, with prior experience at BTG Pactual and ERPFlex, and has developed multiple industry-sponsored AI and machine learning projects during his studies.
Email: thomas.barton@sou.inteli.edu.br

Author 2: Cristiane De Andrade Coutinho
She is a Computer Science undergraduate at Inteli, founder of Inteli Research, and Technology Governance Project Manager Intern at BTG Pactual, passionate about applying artificial intelligence and engineering to generate positive social impact.
Email: cristiane.coutinho@sou.inteli.edu.br

Author 3: Rodolfo Riyoei Goya
He holds B.S. and M.S. degrees in Electrical Engineering and a B.S. in Physics from the University of São Paulo, with extensive experience in networking, systems architecture, and security, and is currently a faculty member at INTELI.
Email: rodolfo.goya@prof.inteli.edu.br

Author 4: Raphael Garcia Moreira
He is Head of AI & Innovation at Eyna, holds MSc and DSc degrees in Electrical Engineering from USP, and specializes in developing privacy-preserving edge AI systems, leading R&D projects, and translating prototypes into production.
Email: raphael.moreira@prof.inteli.edu.br

Corresponding Author: Raphael Moreira

John Doe is the chief investment officer of Acme Investment Management, Inc. Jane Smith is a senior economist at Global Markets Research Ltd.

Algorithm for Adjusting Implied Volatility Curves in the Options Market

September, 2025

Abstract:
Implied volatility is a key metric in the options market, used to align theoretical and observed prices, reflecting market participants' expectations about the future variability of the underlying asset. Although classical models such as Black-Scholes-Merton are widely adopted, they exhibit limitations in capturing the nonlinear structure of volatility, prompting the use of parametric adjustments to account for features such as the volatility "smile" or "smirk". In institutional settings, automated methods are employed to identify key parameters — skew , kurtosis, atm, call wing , and put wing — but these often require manual intervention to correct inaccuracies. This paper proposes an improved algorithm for the automatic fitting of implied volatility curves, aiming to reduce or eliminate the need for manual adjustments. Based on historical data from expert-driven fittings, we developed an algorithm that enhances the accuracy of the automated process, using the root mean square error (RMSE) per ticker as the evaluation metric. The results show a significant improvement over the original algorithm, providing greater consistency and quality in automated fits, and marking an important step toward full automation. The results show a significant improvement over the original algorithm, with 85.71% of tickers performing better, providing greater consistency and quality in automated adjustments and marking an important step towards full automation of volatility surface calibration in the options market.

Keywords: Finance, Risk Management, Asset Allocation

Algorithm for Adjusting Implied Volatility Curves in the Options Market

1 Introduction

The options market constitutes an important component of the modern financial system, enabling investors to take positions on the expected price movements of underlying assets without necessarily acquiring those assets directly [1]. Options are derivative contracts traded both on organized exchanges and in the over-the-counter (OTC) market, and are primarily divided into two types: call options, which grant the holder the right to purchase an asset at a predetermined price by a specified date, and put options, which confer the right to sell the asset under the same conditions. The key parameters of these contracts are the *strike price* and the *maturity date* [2].

Although their prices are linked to the value of the underlying asset through arbitrage mechanisms, options are not perfectly redundant with respect to such assets [3]. This is due to the fact that market volatility, a crucial factor in option pricing, is unknown and time-varying. In this context, options partially reflect a speculative market on volatility itself, known as *Implied Volatility* [4].

Therefore, beyond functioning as instruments for hedging and leverage, options play a fundamental role in shaping expectations and managing risk [4], establishing themselves as essential components of contemporary financial markets.

Implied volatility (IV) is a central metric in the options market, representing the volatility that, when input into a pricing model (typically the Black-Scholes-Merton (BSM) model), equates the theoretical price to the observed market price [5]. It reflects the market's forward-looking expectations about the future variability of the underlying asset and plays a key role in both pricing and risk management. Volatility, in general, measures the degree of randomness or uncertainty associated with changes in an asset's value. High volatility indicates that an asset's price can swing widely over a short period, moving significantly in either direction. In contrast, low volatility suggests more stable and less dramatic price movements, implying reduced uncertainty in the asset's behavior [6].

The Black-Scholes-Merton (BSM) model is one of the most foundational and widely used frameworks in option pricing, serving as a theoretical basis for various financial markets [7]. This model proposes an analytical structure that allows the fair price of a European option to be calculated based on well-defined parameters, such as time to maturity, the risk-free interest rate, the price of the underlying asset, its volatility, and the strike price of the option. Moreover, its applications extend to risk management, stochastic volatility modeling, and transaction cost analysis [7].

An important generalization of the traditional BSM model emerged through the introduc-

tion of time-varying coefficients — specifically, deterministic and differentiable functions — allowing the model to better align with market expectations regarding the behavior of underlying assets [8]. This extension provides greater flexibility to the original model, enabling it to incorporate nonlinear dynamics and changes in market conditions. Through a general transformation, the authors show that the value of a European call option can be expressed as the product of the price obtained using the classical BSM model, the ratio of the underlying stock prices, and a generalized discount factor, thereby enhancing the model's predictive power and adaptability in real-world scenarios [8].

With the advancement of stochastic models and a deeper understanding of the nonlinear dynamics of volatility, there has been growing interest in more robust methods for fitting the implied volatility surface [9]. Although the BSM model is still widely used as a foundation, it has well-known limitations, particularly in directly modeling implied volatility (IV), which leads practitioners to rely on empirically adjusted parameters to capture features such as the volatility "smile" or "smirk" [10].

Several empirical studies have demonstrated that the log-returns of equity assets exhibit stylized features that challenge the assumptions of traditional pricing models. Among these features are the non-normality of return distributions, evidenced by the presence of skewness and excess kurtosis, the stochastic evolution of volatility [11] over time, and the so-called leverage effect, which refers to the negative correlation between returns and their variance. These properties have significant implications for risk modeling and asset pricing, highlighting the need for more robust approaches that better capture the behavior observed in financial markets [12].

Accurately modeling and interpreting these shapes is essential for predicting asset and market volatility. Since implied volatility is derived from observed option prices and is theoretically linked one-to-one with expected future volatility when all other parameters are known, it plays a central role in risk management and derivative pricing. Moreover, under efficient market assumptions and correct model specifications, implied volatility is expected to incorporate all relevant information for forecasting future volatility, surpassing the predictive power of other variables [13].

Several academic studies, such as [14, 15] and [16], have explored alternative approaches to volatility estimation and option pricing, aiming to overcome the limitations of traditional models. A notable example is the work by [17], which proposes a probabilistic approach to estimate the parameters of an option pricing model based on observed market prices. The methodology relies on a stochastic optimization algorithm that generates random samples from the set of global minima of the in-sample pricing error, allowing for the existence of multiple optimal solutions. This approach offers greater flexibility and robustness in estimation, particularly in scenarios characterized by high complexity and uncertainty in market data.

Additionally, other relevant contributions include the use of statistical algorithms and machine learning techniques to enhance volatility forecasting. The study presented in [18] evaluates an ARCH model selection algorithm based on the standardized prediction error criterion (SPEC), demonstrating that appropriate model selection can significantly improve the accuracy of volatility

forecasts in simulated option markets. Complementarily, [19] proposes a data-driven approach using Artificial Neural Networks (ANNs) [20] to price financial options and compute implied volatilities, with the aim of accelerating traditional numerical methods. This proposal highlights the potential of data-driven techniques in handling complex market structures, reinforcing the growing trend of integrating computational methods with financial theory.

In institutional settings, the calibration of the implied volatility curve is often performed using parametric models that aim to identify a set of parameters capable of reproducing the observed curve shape. However, these models are not always sufficiently accurate, requiring human intervention to correct parameter values after automated fitting [21].The algorithm used as a reference for the research seeks to identify these five parameters but still produces results that require manual adjustments by traders or strategists, thereby reducing operational efficiency and increasing response time to market changes.

This paper presents the development of an improved algorithm for fitting implied volatility curves in the options market, focusing on the calibration of the five essential parameters: skewness, kurtosis, at-the-money (ATM), call wing and put wing. The objective of this work is to reduce or eliminate the need for human intervention, increasing the accuracy of the automated curve fitting process. To this end, we receive data containing manually calibrated fits by experts on multiple sets of implied volatility strikes. Based on this data, we develop a new algorithm that replicates and refines human behavior, using the root mean square error (RMSE) per ticker as a comparison metric.

2 Methodology

The primary objective of this work was to develop a robust algorithm for fitting implied volatility surfaces in the options market.The methodology adopted centers on the minimization of an objective function, which incorporates various factors and attributes derived from observed market data, with the goal of accurately and reliably predicting and interpolating volatilities.

2.1 Input data

The algorithm operates using a set of options market information, as presented in Exhibit 1. It requires both the data collection date and the options' due date, which together define the time remaining until expiration. Additionally, the algorithm considers a list of pairs containing each option's strike price and its corresponding volatility.

The current value of the underlying asset (spot) is relevant, as many data adjustments and categorizations depend on the strikes' proximity to this value. A scaling factor is employed to refine the strike fitting, and a set of weights is applied within the Cost Function to assign varying importance to certain data categories. Finally, the number of business days between the collection

Exhibit 1: Options Market Data

t_c	Collection date
t_v	Due date
s	Spot
e	Volatility adjustment factor
W	Cost Function weights
du	Business days between t_c and t_v
K	Strike
V	Volatility
$\mathrm{I} = \{(K_1, V_1), \ldots, (K_n, V_n)\}$	Strike and Volatility

and expiration dates is taken into account to contextualize the data's temporality.

The dataset considered for the analysis consisted of implied volatility curves associated with 56 assets traded on the Brazil Stock Exchange and Over-the-Counter Market (B3) [22] between June 2021 and June 2024, as detailed in Exhibit 2. Each asset was represented by options with different maturities and strike prices. The data were preprocessed to remove inconsistencies and ensure the quality and reliability of the analysis.

2.2 Average of duplicate strikes

The first step of the algorithm is to ensure data quality and consistency. In the options market, it is common to find multiple records for the same strike price but with different volatilities. For categorization, the algorithm identifies duplicate strikes and calculates the average of their corresponding volatilities. These data points are stored and used to adjust the curve such that it passes as closely as possible between the two volatility points, functioning similarly to a tunnel between the values.

For identifying the set of pairs with similar strikes, the following is considered:

$$\begin{aligned} I_p = \Bigg\{ & \left(K_i, \frac{V_i + V_j}{2} \right) \\ & \mid (i,j) \in \{1, \ldots, n\}^2 \\ & \mid i < j, \quad K_i = K_j \} \end{aligned} \tag{1}$$

2.3 Identifying Outliers with RANSAC

The robustness of volatility curve fitting is of great importance, as market data can contain noise or outliers. To address this, the Random Sample Consensus (RANSAC) algorithm is employed. This method is chosen for its ability to fit a model to a dataset that contains a significant number

of outliers [23]. It has applications in automated image analysis, where interpretation is based on data provided by error-prone feature detectors [24], as well as in estimating scale and translation parameters in point cloud registration tasks [25], among others.

In essence, RANSAC randomly selects a small subset of data (the "inliers"), fits a model (in this case, a 3rd-degree polynomial, chosen as it generally provides a good compromise between capturing the typical curvature of a volatility smile and avoiding overfitting to noise for the density of data points usually observed in option chains) to these points, and then verifies how many other points from the original set fit well within that model, given a certain error threshold. This process is repeated multiple times, and the model that best fits the largest number of inliers is selected. This allows the algorithm to disregard outliers, which would deviate significantly from any model consistent with the majority of the data. The result is a set of reliable inliers, which will be used in the optimization steps, ensuring that the final curve is not distorted by erroneous data.

In the reference technique, the Isolation Forest was used—an unsupervised machine learning algorithm designed to detect anomalies in datasets [26, 27, 28]. In the present work, a polynomial transformation was applied to the strike values. For each $K_i \in VK$, the transformation is given by:

$$\Phi_3(K_i) = \begin{bmatrix} 1 & K_i & K_i^2 & K_i^3 \end{bmatrix} \tag{2}$$

This transformation selects each K_i value and creates a four-element vector related to it. Subsequently, the model was fitted to the data using RANSAC regression. The fitted curve is described by the function:

$$\hat{V}_i = f(K') = \beta_0 + \beta_1 K_i + \beta_2 K_i^2 + \beta_3 K_i^3 \tag{3}$$

where $\hat{V}_i$ represents the predicted volatilities, and β_0, β_1, β_2, β_3 are the polynomial coefficients that minimize the error between the observed volatilities V_i and the predicted volatilities $\hat{V}_i$.

In this regard, RANSAC determines two sets of points: inliers, which are points close to the fitted function, and outliers, which are points distant from the fitted function.

The difference between the observed volatility V_i and the fitted volatility $\hat{V}_i$ is compared against an error threshold ϵ. This threshold defines the maximum deviation a point can exhibit from the fitted curve while still being considered valid (an inlier). If the difference between the actual and predicted value exceeds ϵ, the point is classified as an outlier.

Formally, we define:

$$I_{\text{l}} = \left\{ (K_i, V_i) \in I \;\middle|\; |V_i - \hat{V}_i| \leq \epsilon \right\}. \tag{4}$$

The error threshold ϵ was set to 0.20, representing an absolute volatility difference. This value was determined after conducting multiple tests with thresholds ranging from 0.10 to 1.00, where 0.20 consistently yielded the most robust results.

2.4 Data Categorization and Adjustment for Optimization

To ensure that the volatility curve adjustment reflects the characteristics observed in the market, the data is categorized and adjusted. This categorization is based on market patterns where certain points are more relevant to the shape of the curve. For example, strikes close to the spot price (current asset price) and the so-called "tunnels" (points between strikes of the same value) are considered of greater importance.

Based on this logic, the algorithm organizes the information into four groups: inners, referring to strikes close to the spot price; outers, which represent strikes farther from the spot; filtered, corresponding to tunnels in the vicinity of the spot price, regarded as highly relevant; and valid, which comprise tunnels located at greater distances. This classification enables the cost function to apply differentiated priorities to each group, in accordance with the predefined weights.

Thus, let:

- I_K: be the set of indices of I_l
- I_P: be the set of indices of I_p
- $\epsilon_1 = 0.2$: A relative threshold defining the 'near-the-money' region for inlier strikes as those within $\pm 20\%$ of the spot price s.
- $\epsilon_2 = 0.1$: A relative threshold defining important tunnels as those whose strikes are within $\pm 10\%$ of the total inlier strike range from the spot price s.
- V_s: Volatility of the strike closest to the spot (s)
-

$$V_i^{(\text{t})} = \left| V_i^{(\text{l})} - \left((V_i^{(\text{l})} - V_s)(1 - e) + V_s \right) \right|, \tag{5}$$
$$i \in I_K, \quad V_i^{(\text{l})} \in \{ V_j \mid (K_j, V_j) \in I_l \}$$

There are:

$$
\begin{aligned}
I_{\text{inners}} = \{\, & K_i^{(\text{l})} \geq (1-\epsilon_1)s \\
& \wedge\ K_i^{(\text{l})} \leq (1+\epsilon_1)s \,\}, \\
& i \in I_K, \quad K_i^{(\text{l})} \in \{(K_j, V_j) \in I^{(\text{l})}\}
\end{aligned}
\tag{6}
$$

$$
\begin{aligned}
I_{\text{outers}} = \{\, & K_i^{(\text{l})} < (1-\epsilon_1)s \\
& \vee\ K_i^{(\text{l})} > (1+\epsilon_1)s \,\}, \\
& i \in I_K, \quad K_i^{(\text{l})} \in \{(K_j, V_j) \in I^{(\text{l})}\}
\end{aligned}
\tag{7}
$$

$$
\begin{aligned}
I_{\text{filtered}} = & \left\{ \left| K_i^{(\text{p})} - s \right| \leq \epsilon_2 \cdot \left(K_{\max}^{(\text{l})} - K_{\min}^{(\text{l})} \right) \right\}, \\
& i \in I_P, \quad K_i^{(\text{p})} \in \{K_j \mid (K_j, V_j) \in I^{(\text{p})}\}
\end{aligned}
\tag{8}
$$

$$
\begin{aligned}
I_{\text{valid}} = & \left\{ \left| K_i^{(\text{p})} - s \right| > \epsilon_2 \cdot \left(K_{\max}^{(\text{l})} - K_{\min}^{(\text{l})} \right) \right\}, \\
& i \in I_P, \quad K_i^{(\text{p})} \in \{K_j \mid (K_j, V_j) \in I^{(\text{p})}\}
\end{aligned}
\tag{9}
$$

Furthermore, the inlier strikes are subjected to a linear adjustment, given by:

$$
\begin{aligned}
V_{\text{adjusted}} = & \left\{ \begin{array}{ll} V_i^{(\text{l})} - V_i^{(\text{t})}, & \text{if } V_i^{(\text{l})} - V_i^{(\text{t})} \geq V_s \\ V_i^{(\text{l})} + V_i^{(\text{t})}, & \text{otherwise} \end{array} \right\}, \\
& i \in I_K, \quad V_i^{(\text{l})} \in \{V_j \mid (K_j, V_j) \in I\}
\end{aligned}
\tag{10}
$$

This adjustment aims to systematically pull the inlier volatilities $V_i^{(\text{l})}$ towards the spot volatility V_s by applying the previously calculated adjustment magnitude $V_i^{(\text{t})}$. The conditional logic ensures this pull is directional towards V_s:

- If $V_i^{(\text{l})} - V_i^{(\text{t})} \geq V_s$: This implies $V_i^{(\text{l})}$ is sufficiently above V_s such that even after tentatively subtracting $V_i^{(\text{t})}$, it remains above or at V_s. In this case, $V_i^{(\text{t})}$ is subtracted from $V_i^{(\text{l})}$.
- Otherwise: This implies $V_i^{(\text{l})}$ is below V_s, or close enough to V_s that subtracting $V_i^{(\text{t})}$ would cause it to undershoot V_s (or move further away if already below). In this scenario, $V_i^{(\text{t})}$ is added to $V_i^{(\text{l})}$.

The effect is to reduce the deviation from V_s, with the magnitude of $V_i^{(\text{t})}$ being influenced by the

factor e and the initial distance of $V_i^{(1)}$ from V_s. This adjustment is vital for the final curve to better adapt to the volatility structure observed near the spot, a region of typically greatest liquidity and pricing interest.

2.5 Cost Function

The central stage of the algorithm involves the adjustment of the implied volatility surface, achieved through the minimization of a cost function. This function represents the weighted sum of squared errors between the volatilities predicted by the parametric model and the observed market volatilities using the curve fitting function, shown in Algorithm 1. The parametric model employs a set of parameters (such as skew, kurtosis, atm, callwing, and putwing) that define the shape of the volatility curve, enabling it to capture the typical characteristics of the volatility surface.

The cost function is formulated to assign different weights W_j (from the input W vector, see Exhibit 1) to each data category (filtered, valid, inners, outers, and adjusted, respectively), prioritizing the subsets most relevant to the curve's adjustment, as shown in Algorithm 2. The values of W are defined by the user; otherwise, default values are retained. This implies that errors at more significant points (such as tunnels or strikes near the spot) will have a greater impact on the cost function, compelling the algorithm to better conform to these points. The optimization of these parameters is performed using numerical methods, which seek the set of parameters that minimizes this cost function, resulting in the adjusted implied volatility curve that best represents the market data, as shown in Algorithm 3.

In this regard, the function is given by:

$$\theta = (\alpha,\ \beta,\ \gamma,\ \delta,\ \nu) \tag{11}$$

$$\begin{aligned} f(K_i;\theta) = \gamma + 10 \cdot \alpha \cdot w_i \cdot \arctan\left(-\frac{n_i}{w_i}\right) \\ + \beta \cdot \left(10 \cdot w_i \cdot \arctan\left(-\frac{n_i}{w_i}\right)\right)^2 \end{aligned} \tag{12}$$

In which:

$$
\begin{aligned}
\min_{\theta} \quad \mathcal{L}(\theta) = W_1 \cdot & \sum_{i \in I_{\text{filtered}}} \left(f(K_i^{(\text{p})}; \theta) - V_i^{(\text{p})} \right)^2 \\
+ W_2 \cdot & \sum_{i \in I_{\text{valid}}} \left(f(K_i^{(\text{p})}; \theta) - V_i^{(\text{p})} \right)^2 \\
+ W_3 \cdot & \sum_{i \in I_{\text{inners}}} \left(f(K_i^{(\text{l})}; \theta) - V_i^{(\text{l})} \right)^2 \\
+ W_4 \cdot & \sum_{i \in I_{\text{outers}}} \left(f(K_i^{(\text{l})}; \theta) - V_i^{(\text{l})} \right)^2 \\
+ W_5 \cdot & \sum_{i \in I_K} \left(f(K_i^{(\text{l})}; \theta) - V_i^{(\text{adjusted})} \right)^2
\end{aligned}
\tag{13}
$$

The algorithm was implemented in Python, with the support of the `pandas`, `NumPy`, `SciPy`, and `scikit-learn` libraries. The cost function adopted for the adjustment was the Mean Squared Error (MSE), measured between the values adjusted by the computational curve and the values from the provided human fit. The optimization of the curve parameters was performed using the `minimize` method from the `scipy.optimize` library.

Algorithm 1 Vectorized Volatility Fitting Function

Require: `strike`, `skew`, `kurtose`, `atm`, `callwing`, `putwing`, `Spot`, `DU`
Ensure: `volatility`: Vector of calculated volatilities

1: $T \leftarrow \texttt{DU}/252$ ▷ Convert business days to years
2: $ns \leftarrow \ln(\texttt{strike}/\texttt{Spot})/\sqrt{T}$ ▷ Normalized strike (vectorized)
3: $\texttt{wing} \leftarrow \texttt{callwing}$ where $ns \leq 0$, else $\texttt{putwing}$ ▷ Conditional selection (vectorized)
4: $\texttt{arctgwing} \leftarrow \arctan(-ns/\texttt{wing})$ ▷ Element-wise arctangent
5: $\texttt{volatility} \leftarrow \texttt{atm} + (\texttt{skew} \cdot 10 \cdot \texttt{wing} \cdot \texttt{arctgwing}) + \texttt{kurtose} \cdot (10 \cdot \texttt{wing} \cdot \texttt{arctgwing})^2$ ▷ Smile fitting formula
6: $\texttt{volatility}[i] \leftarrow 0$ **where** $\texttt{volatility}[i] < 0$ ▷ Ensure non-negative values
7: **return** `volatility`

2.6 Wilcoxon Test

To evaluate the results and compare the values obtained by the reference technique and the suggested algorithm, the Wilcoxon test was applied. The Wilcoxon test is a non-parametric statistical method widely used to assess the existence of significant differences between two related samples or experimental conditions, especially when the normality assumptions required by parametric tests are not met [29]. This test is based on the ranks of the observed differences between paired values, allowing inferences about the central tendency without assuming a specific data distribution. In general terms, let n and m denote the sample sizes of the two independent random samples, each consisting of independent and identically distributed observations, with independence also

Algorithm 2 Objective Function for Volatility Calibration

Require: *params*, *weights*, strike and volatility sets: *filtered*, *valid*, *inlier*, *adj*, *adjusted*; underlying *Spot*, time to expiry *DU*

Ensure: *total_penalty*: weighted sum of squared errors

$[skew, kurtose, atm, callwing, putwing] \leftarrow params$
$pred\_filtered \leftarrow$ fitting_function(
    filtered_strikes, *skew*, *kurtose*, *atm*, *callwing*, *putwing*, *Spot*, *DU*)
$penalty\_filtered \leftarrow \sum(pred\_filtered - filtered\_volatilities)^2$
$pred\_valid \leftarrow$ fitting_function(*valid_strikes*, ...)
$penalty\_valid \leftarrow \sum(pred\_valid - valid\_volatilities)^2$
$pred\_l \leftarrow$ fitting_function(*l_strikes*, ...)
$penalty\_l \leftarrow \sum(pred\_l - l\_volatilities)^2$
$pred\_adj \leftarrow$ fitting_function(*adj_strikes*, ...)
$penalty\_adj \leftarrow \sum(pred\_adj - adj\_volatilities)^2$
$pred\_adjusted \leftarrow$ fitting_function(*adjusted_strikes*, ...)
$penalty\_adjusted \leftarrow \sum(pred\_adjusted - adjusted\_volatilities)^2$
$total\_penalty \leftarrow$
    $weights_0 \cdot penalty\_filtered + weights_1 \cdot penalty\_valid +$
    $weights_2 \cdot penalty\_l + weights_3 \cdot penalty\_adj +$
    $weights_4 \cdot penalty\_adjusted$
return *total_penalty*

Algorithm 3 Optimization of Volatility Smile Parameters

Require: *objective_function*, *weights*, strike and volatility sets, *Spot*, *DU*

Ensure: *result*: Optimal parameters from the fitting procedure

bounds ← [(-10, ∞), (0.1, ∞), (0, ∞), (0.1, ∞), (0.1, ∞)]
initial_guess ← [0.1, 0.1, 0.1, 0.1, 0.1]
args ← (weights, filtered_strikes, filtered_volatilities, valid_strikes, valid_volatilities,
    inlier_strikes, inlier_volatilities, adj_strikes, adj_volatilities,
    adjusted_strikes, adjusted_volatilities, Spot, DU)
result ← Minimize(*objective_function*, *initial_guess*,
    args=args, *method='SLSQP'*, *bounds=bounds*)

assumed between the two groups [29]. The null hypothesis states that there is no systematic difference between the compared groups, whereas the alternative hypothesis aims to detect a significant shift in their relative locations. The Wilcoxon test is particularly useful in applied settings, such as educational research and experimental studies, where sample sizes may be small or the data may not follow a normal distribution [30].

3 Results and Discussion

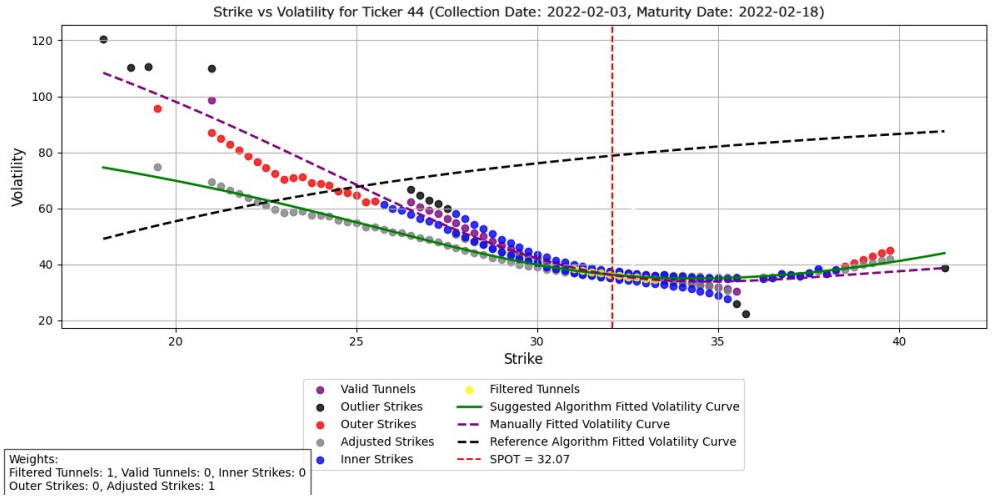
In this section, an evaluation of the performance of the proposed algorithm for fitting implied volatility curves is presented. The analysis is conducted using quantitative metrics (mean squared error – MSE) [31], comparisons with the method previously used by the reference technique, and non-parametric statistical tests (Wilcoxon signed-rank test) [32].

As shown in Exhibit 2, the proposed algorithm achieved superior performance in 85.71% of the tickers. The largest improvement in MSE was observed for Ticker 13, which recorded an improvement of 1E+02%. A similar enhancement was observed for Ticker 25, as well as for Ticker 27, Ticker 45, and Ticker 49.

Although the algorithm generally outperformed the benchmark, a minority of eight tickers exhibited worse performance. The lowest result was observed for ticker Ticker 54, with a negative improvement of -1.50E+06. Other tickers with inferior performance included Ticker 51, Ticker 2, Ticker 44, Ticker 37, Ticker 31, and Ticker 31.

These cases exhibiting highly negative improvements (e.g., Ticker 54, Ticker 51, Ticker 2, Ticker 44) contain outliers that may indicate specific weaknesses in the algorithm, depending on the asset's volatility, liquidity, or time-series characteristics.

Exhibit 1: Comparison of the results of the algorithms for the Ticker 44.



The disparity between the algorithms can also be more clearly observed when viewed in

Exhibit 2: Comparison of the results of the volatility curve algorithms

	Relative		Improvements	
Ticker	**STD**	**MSE**	**STD %**	**MSE %**
Ticker 1	1.1E-01	3.9E-02	8.9E+01	9.6E+01
Ticker 2	1.7E+01	1.0E+01	-1.6E+03	-9.4E+02
Ticker 3	9.1E-03	1.5E-02	9.9E+01	9.8E+01
Ticker 4	1.4E-01	9.4E-02	8.6E+01	9.1E+01
Ticker 5	7.1E-01	1.9E-01	2.9E+01	8.1E+01
Ticker 6	1.8E-02	1.5E-02	9.8E+01	9.8E+01
Ticker 7	2.0E-02	8.1E-03	9.8E+01	9.9E+01
Ticker 8	9.2E-02	5.1E-02	9.1E+01	9.5E+01
Ticker 9	1.7E-02	8.1E-03	9.8E+01	9.9E+01
Ticker 10	8.8E-01	1.1E+00	1.2E+01	-1.2E+01
Ticker 11	8.0E-01	5.9E-02	2.0E+01	9.4E+01
Ticker 12	3.4E-02	9.6E-03	9.7E+01	9.9E+01
Ticker 13	2.2E-03	1.6E-03	1.0E+02	1.0E+02
Ticker 14	4.8E-01	2.8E-01	5.2E+01	7.2E+01
Ticker 15	3.0E-01	1.1E-01	7.0E+01	8.9E+01
Ticker 16	3.7E-03	5.9E-03	1.0E+02	9.9E+01
Ticker 17	7.3E-02	2.0E-02	9.3E+01	9.8E+01
Ticker 18	1.2E-02	6.4E-03	9.9E+01	9.9E+01
Ticker 19	4.1E-01	2.2E-01	5.9E+01	7.8E+01
Ticker 20	1.2E-02	1.2E-02	9.9E+01	9.9E+01
Ticker 21	5.9E-01	3.1E-01	4.1E+01	6.9E+01
Ticker 22	9.8E-02	4.8E-02	9.0E+01	9.5E+01
Ticker 23	1.1E-01	5.8E-02	8.9E+01	9.4E+01
Ticker 24	2.7E-02	1.5E-02	9.7E+01	9.9E+01
Ticker 25	1.3E-04	2.0E-04	1.0E+02	1.0E+02
Ticker 26	4.2E-03	5.7E-03	1.0E+02	9.9E+01
Ticker 27	1.1E-03	1.8E-03	1.0E+02	1.0E+02
Ticker 28	6.1E-02	1.9E-02	9.4E+01	9.8E+01
Ticker 29	2.2E-01	5.6E-02	7.8E+01	9.4E+01
Ticker 30	2.5E-02	1.2E-02	9.8E+01	9.9E+01
Ticker 31	3.4E+00	1.5E+00	-2.4E+02	-5.3E+01
Ticker 32	2.5E+00	8.2E-01	-1.5E+02	1.8E+01
Ticker 33	8.5E-02	2.1E-02	9.2E+01	9.8E+01
Ticker 34	5.4E-02	1.7E-02	9.5E+01	9.8E+01
Ticker 35	2.2E-01	9.0E-02	7.8E+01	9.1E+01
Ticker 36	1.2E-01	6.1E-02	8.8E+01	9.4E+01
Ticker 37	1.2E+01	2.2E+00	-1.1E+03	-1.2E+02
Ticker 38	6.0E-01	1.5E-01	4.0E+01	8.5E+01
Ticker 39	2.2E-01	7.5E-02	7.8E+01	9.3E+01
Ticker 40	1.4E-02	1.4E-02	9.9E+01	9.9E+01
Ticker 41	4.9E-01	1.3E-01	5.1E+01	8.7E+01
Ticker 42	2.7E-02	1.1E-02	9.7E+01	9.9E+01
Ticker 43	2.2E-02	6.7E-03	9.8E+01	9.9E+01
Ticker 44	4.2E+01	3.7E+00	-4.1E+03	-2.7E+02
Ticker 45	4.7E-04	7.9E-04	1.0E+02	1.0E+02
Ticker 46	2.1E-01	5.1E-02	7.9E+01	9.5E+01
Ticker 47	6.8E-02	1.6E-02	9.3E+01	9.8E+01
Ticker 48	1.5E-02	7.9E-03	9.8E+01	9.9E+01
Ticker 49	9.5E-04	1.8E-03	1.0E+02	1.0E+02
Ticker 50	6.8E-02	3.0E-02	9.3E+01	9.7E+01
Ticker 51	1.6E+04	1.7E+03	-1.6E+06	-1.7E+05
Ticker 52	0	1.1E+00	0	-9.2E+00
Ticker 53	4.6E-01	2.2E-01	5.4E+01	7.8E+01
Ticker 54	1.7E+05	1.5E+04	-1.7E+07	-1.5E+06
Ticker 55	1.4E-02	6.9E-03	9.9E+01	9.9E+01
Ticker 56	5.6E-02	1.4E-02	9.4E+01	9.9E+01

Exhibit 1, where the curve of the proposed algorithm shows greater similarity to the ideal curve than the curve generated by the reference technique. It is important to note that this is the volatility curve for a single ticker (Ticker 44), and therefore the result cannot be generalized to other tickers. The graph is provided merely as an example of the differences that exist between the outputs of the two algorithms.

It is noteworthy that volatility itself is widely studied in the financial literature, with some works indicating that even volatility behaves as a significant risk factor that affects returns and rates of volatility of index options [33], as well as several approaches to determination of implicit volatility for. Options [34]. It should also be considered the period in which these tickets have been removed, the behavior and adaptation of the algorithm may have different performance, depending on the context and characteristics of the tickets, as the author [35] pointed to comparison between the nineteenth-century market and 21st century.

As previously mentioned, the Wilcoxon signed-rank test was applied to assess whether the difference in errors between the two methods is statistically significant. The test results yielded a p-value of 8E-6 and a test statistic of 384, confirming a statistically significant difference between the algorithms.

Thus, the improvements observed through the application of the proposed algorithm not only enhance the accuracy of fitting implied volatility curves but also contribute significantly to operational efficiency. These gains translate into reduced costs associated with manual processing and human intervention in curve adjustments, as well as a substantial increase in the automation of the options pricing process. As a result, the proposed model demonstrates to be a viable and scalable solution for environments requiring frequent and rapid updates of derivative instrument prices, such as the options market traded on B3.

Several studies have explored volatility modeling and its impact on derivatives pricing. A flexible framework for jointly modeling an index and term swap rates written on it was proposed in [36], allowing for consistent pricing of volatility derivatives and options while capturing empirical features such as jumps in volatility and returns. In contrast, traditional models often rely on historical or realized volatility to forecast future levels [37]. That study uses implied volatility from S&P 100 index options to show that volatility changes can be predicted, although this predictability does not lead to abnormal returns when transaction costs are considered — suggesting compatibility with market efficiency.

Other works have focused on the relationship between implied volatility, moneyness, and option pricing. Reference [38] investigates this link using WTI crude oil options through a two-stage regression approach, revealing how volatility varies across strike prices and maturities. These findings highlight the importance of modeling volatility surfaces accurately, which is directly relevant to improving calibration and performance evaluation methods — objectives also central to our proposed algorithm for fitting implied volatility curves.

4 Conclusion

In conclusion, the proposed algorithm demonstrates superior performance in fitting implied volatility curves when compared to the method previously used by the reference technique. The quantitative analysis, based on the mean squared error relative (MSE), revealed that 85.71% of the evaluated tickers (48 out of 56) achieved lower average errors with the new approach. Despite the high standard deviation observed across assets—which reflects the inherent heterogeneity of the underlying instruments, diverse market conditions, and the presence of notable outliers such as Ticker 54, Ticker 51, Ticker 2, and Ticker 44—the overall results indicate a statistically and practically significant improvement in fitting accuracy.

This improvement is further corroborated by the Wilcoxon signed-rank test, which confirmed that the difference in errors between the two methods is statistically significant (p-value = 7e-4). Importantly, the algorithm's enhanced precision is not limited to a narrow subset of assets; it consistently outperforms the benchmark across a broad spectrum of tickers, including those with varying levels of liquidity, volatility regimes, and maturity structures.

Beyond accuracy, the proposed algorithm contributes meaningfully to operational efficiency. By producing more reliable and stable volatility curves, it reduces the need for manual adjustments and expert intervention, thereby enabling a higher degree of automation in the options pricing pipeline. This is particularly valuable in high-frequency trading environments and institutional settings—such as the B3 options market—where timely and accurate derivative valuations are critical for risk management, hedging, and pricing strategies.

As a future extension of this work, we suggest implementing a minimization model tailored to each individual ticker. By optimizing weights, scaling factors, and other relevant parameters at the ticker level—potentially informed by asset-specific features such as trading volume, bid-ask spreads, or historical volatility clustering—it may be possible to further reduce the MSE and enhance the robustness of the curve-fitting process. Such an adaptive calibration strategy could mitigate the impact of outliers and better accommodate assets with idiosyncratic volatility dynamics or limited market depth. Additionally, integrating machine learning techniques to dynamically adjust model hyperparameters based on real-time market data represents a promising avenue for future research, potentially bridging the gap between parametric flexibility and empirical fidelity in implied volatility modeling.

Acknowledgment

The authors would like to express their gratitude to the partner bank for the collaboration, availability of data, and continuous support throughout the project. Special thanks are also extended to Professor Henrique Paiva and Professor Crishna Irion for their valuable guidance, insights, and academic supervision, which were fundamental to the successful development of this work.

References

[1] M. Ansbacher, *The New Options Market*. John Wiley & Sons, 2000.

[2] J. Hull, S. Treepongkaruna, D. Colwell, R. Heaney, and D. Pitt, *Fundamentals of Futures and Options Markets*. Pearson Higher Education AU, 2013.

[3] D. J. Skinner, "Options markets and stock return volatility," *Journal of Financial Economics*, vol. 23, no. 1, pp. 61–78, 1989.

[4] J. Stein, "Overreactions in the options market," *The Journal of Finance*, vol. 44, no. 4, pp. 1011–1023, 1989.

[5] T. Yue, S. A. Gehricke, J. E. Zhang, and Z. Pan, "The implied volatility smirk in the chinese equity options market," *Pacific-Basin Finance Journal*, vol. 69, p. 101624, 2021.

[6] B. Nabubie and S. Wang, "Numerical techniques for determining implied volatility in option pricing," *Journal of Computational and Applied Mathematics*, vol. 422, p. 114913, 2023.

[7] M. Qayyum, E. Ahmad, F. M. Tawfiq, Z. Salleh, S. T. Saeed, and M. Inc, "Series form solutions of time–space fractional black–scholes model via extended he-aboodh algorithm," *Alexandria Engineering Journal*, vol. 109, pp. 83–88, 2024.

[8] P. Morales-Bañuelos, N. Muriel, and G. Fernández-Anaya, "A modified black–scholes–merton model for option pricing," *Mathematics*, vol. 10, no. 9, p. 1492, 2022.

[9] C. S. Jones, "The dynamics of stochastic volatility: evidence from underlying and options markets," *Journal of econometrics*, vol. 116, no. 1-2, pp. 181–224, 2003.

[10] S. Mixon, "Option markets and implied volatility: Past versus present," *Journal of Financial Economics*, vol. 94, no. 2, pp. 171–191, 2009.

[11] N. Shephard, *Stochastic volatility: selected readings*. OUP Oxford, 2005.

[12] L. Ballotta and G. Rayée, "Smiles & smirks: Volatility and leverage by jumps," *European Journal of Operational Research*, vol. 298, no. 3, pp. 1145–1161, 2022.

[13] D. Liu *et al.*, "Implied volatility forecast and option trading strategy," *International Review of Economics & Finance*, vol. 71, pp. 943–954, 2021.

[14] A. Gupta, C. Reisinger, and A. Whitley, "Model uncertainty and its impact on derivative pricing," *Risk*, 2011.

[15] X. J. He and S. Lin, "A stochastic liquidity risk model with stochastic volatility and its applications to option pricing," *Stochastic Models*, vol. 41, no. 3, pp. 273–292, 2024.

[16] X. J. He and W. Chen, "A closed-form pricing formula for european options under a new stochastic volatility model with a stochastic long-term mean," *Mathematics and Financial Economics*, vol. 15, pp. 381–396, 2021.

[17] R. Cont and S. Ben Hamida, "Recovering volatility from option prices by evolutionary optimization," *SSRN*, 2004.

[18] E. Xekalaki and S. Degiannakis, "Evaluating volatility forecasts in option pricing in the context of a simulated options market," *Computational statistics & data analysis*, vol. 49, no. 2, pp. 611–629, 2005.

[19] S. Liu, C. W. Oosterlee, and S. M. Bohte, "Pricing options and computing implied volatilities using neural networks," *Risks*, vol. 7, no. 1, p. 16, 2019.

[20] Y. Li and W. Ma, "Applications of artificial neural networks in financial economics: a survey," in *2010 International symposium on computational intelligence and design*, vol. 1, pp. 211–214, IEEE, 2010.

[21] P. Carr and L. Wu, "Analyzing volatility risk and risk premium in option contracts: A new theory," *Journal of Financial Economics*, vol. 120, no. 1, pp. 1–20, 2016.

[22] Brazil Stock Exchange and Over-the-Counter Market (B3), "B3 social: Annual report 2023." https://www.b3.com.br/data/files/E3/60/F7/72/B6CC09105FE89209AC094EA8/B3-SOCIAL-RA-2023$_{A}cessivel.pdf, 2024. Accessed : 2025 - 06 - 25.$

[23] H. Cantzler, "Random sample consensus (ransac)," tech. rep., Institute for Perception, Action and Behaviour, Division of Informatics, University of Edinburgh, 1981.

[24] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, pp. 381–395, June 1981.

[25] J. Li, Q. Hu, and M. Ai, "Point cloud registration based on one-point ransac and scale-annealing biweight estimation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 11, pp. 9716–9729, 2021.

[26] H. Xu, G. Pang, Y. Wang, and Y. Wang, "Deep isolation forest for anomaly detection," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 12, pp. 12591–12604, 2023.

[27] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation forest," in *2008 eighth ieee international conference on data mining*, pp. 413–422, IEEE, 2008.

[28] S. Hariri, M. C. Kind, and R. J. Brunner, "Extended isolation forest," *IEEE transactions on knowledge and data engineering*, vol. 33, no. 4, pp. 1479–1489, 2019.

[29] N. I. S. Indah and F. Ahmad, "Exploring the wilcoxon test in science education: A literature review of empirical research," *Indonesian Journal of Educational Science (IJES)*, vol. 7, no. 2, pp. 158–169, 2025.

[30] M. Neuhäuser, "Wilcoxon–mann–whitney test," in *International Encyclopedia of Statistical Science* (M. Lovrić, ed.), pp. 1656–1658, Berlin, Heidelberg: Springer, 2011.

[31] M. D. Schluchter, "Mean square error," *Encyclopedia of Biostatistics*, vol. 5, 2005.

[32] E. Fix and J. Hodges Jr, "Significance probabilities of the wilcoxon test," *The Annals of Mathematical Statistics*, pp. 301–312, 1955.

[33] D. Huang, C. Schlag, I. Shaliastovich, and J. Thimme, "Volatility-of-volatility risk," *Journal of Financial and Quantitative Analysis*, vol. 54, no. 6, p. 2423–2452, 2019.

[34] G. Orlando and G. Taglialatela, "A review on implied volatility calculation," *Journal of Computational and Applied Mathematics*, vol. 320, pp. 202–220, 2017.

[35] S. Mixon, "Option markets and implied volatility: Past versus present," *Journal of Financial Economics*, vol. 94, no. 2, pp. 171–191, 2009.

[36] C. R. Harvey and R. E. Whaley, "Market volatility prediction and the efficiency of the s & p 100 index option market," *Journal of Financial Economics*, vol. 31, no. 1, pp. 43–73, 1992.

[37] P. A. Abken and S. Nandi, "Options and volatility," *Economic Review*, vol. 81, no. 3-6, p. 21, 1996.

[38] V. Soini and S. Lorentzen, "Option prices and implied volatility in the crude oil market," *Energy Economics*, vol. 83, pp. 515–539, 2019.