Parameter calibration of stochastic volatility
Heston’s model: Constrained optimization vs.
differential evolution

Ambrosio Ortiz-Ramírez, Francisco Venegas-Martínez,
María Teresa Verónica Martínez-Palacios*

Instituto Politécnico Nacional, México

Received November 20, 2019; accepted May 13, 2021
Available online May 19, 2021

Abstract

This paper calibrates through loss functions the parameters of Heston’s stochastic volatility model by using two different methods: minimizing a nonlinear objective function (a loss function) with constraints on the values of the parameter and using a differential evolution algorithm. Both methods are applied to implied volatilities on the Mexican Stock Exchange Index with four maturities and twenty-eight strike prices. The selection criterion for the parameters is minimizing the value of the mean square error of the implied volatility. The first method has a better performance with less error and time. However, empirical results show that for both methods the adjustment of implied volatilities is better for options with long-term maturities than for short-term maturities.

JEL Code: G13, B16, C21
Keywords: Contingent pricing; Stochastic volatility; Implied volatility; Differential evolution

* Corresponding author.
E-mail address: amortiz@ipn.mx (A. Ortiz-Ramírez), fvenegas1111@yahoo.com.mx (Francisco Venegas-Martínez), mmartinezpa@ipn.mx (M. T. V. Martínez-Palacios).
Peer Review under the responsibility of Universidad Nacional Autónoma de México.

http://dx.doi.org/10.22201/fca.24488410e.2022.2789
0186- 1042©2019 Universidad Nacional Autónoma de México, Facultad de Contaduría y Administración. This is an open access article under the CC BY-NC-SA (https://creativecommons.org/licenses/by-nc-sa/4.0/)
Introduction

In the last decades, many stochastic volatility models (SVMs) for option pricing have arisen as an alternative (or generalization) to the standard Black-Scholes-Merton (BSM) model (Black and Scholes, 1973; and Merton, 1973). Among the main limitations of the BSM model is the assumption of constant volatility. Stochastic volatility models undertake this problem under the assumption that prices and their volatility are stochastic processes affected by several risk factors. In this regard, SVMs handle more consistent prices. Unlike the BSM model, the SVMs considers stylized facts, such as implied volatility smile and skewness, as observed in derivative markets. The latter handle the leverage effect, i.e., volatility rises after price declines. That is, the asset price and its volatility are negatively correlated. The best known SVMs are those from Hull and White (1987), Stein and Stein (1991), Heston (1993), and Schöbel and Zhu (1999). In particular, Hull and White’s (1987) approach expand the asset price with a Taylor series function up to terms of third order, and Heston’s (1993) model provides a closed formula that requires numeric methods.¹

A main issue of option pricing models is their capability of generating volatility structures that are consistent with the expiration date and the strike price in order to capture the behavior of the underlying asset return distribution. There are three important features to be considered once a model is proposed: tail risk, option pricing, and implied volatility. Tail risk is empirically defined as the probability that the underlying asset return will move more than three standard deviations from its mean. The BSM model underestimates tail risk since the forecasted probability for extreme variations is low. This is because in the BSM model the tails of the asset return distribution are driven by a normal distribution. Thus, in this context the models of Hull and White (1987), Stein and Stein (1991), and Heston (1993) exhibit higher performance than the BSM model allowing estimating the probabilities for unexpected falls in the asset price (Gulisashvili and Stein, 2010).

Volatility is an unobservable variable that measures the relative changes in an asset price. Often the volatility is also seen as the instantaneous standard deviation of the asset returns. It important to point out that through periods of low volatility, asset prices do not change very much; however, in periods of high volatility, significant variations are shown for the asset prices. In the BSM model, volatility is constant and positive until the expiration date. In contrast, a SVM, volatility is described by a non-negative stochastic process. For example, in Hull and White’s (1987) model, volatility dynamics follows a

¹ More work dealing with calibrating SVM can be found in: Venegas-Martínez (2005), (2010), Ortiz-Ramírez, Venegas-Martínez and López-Herrera (2011), and Ortiz-Ramírez, Venegas-Martínez and Durán-Bustamante (2014).
geometric Brownian motion. In Stein and Stein’s (1991) model, volatility follows an Ornstein-Uhlenbeck process or the absolute value. While in Heston’s (1993) model, volatility is driven by the square root of the Cox-Ingersoll-Ross’ (1985) process.

The term “implied volatility” appeared for the first time in the work of Latané and Rendleman (1976), called “implied standard deviation.” These authors studied the standard deviations of returns on assets implied in observed call options prices when an investor values options prices under the standard BSM model. When implied volatility is obtained with options price data, and the strike price and expiration date are plot, the implied volatility surface is not flat, and changes along with the strike price and expiration date. This is what is known as implied volatility smile.

In relation to the estimation of the parameters of stochastic volatility, there are three main statistical techniques: maximum likelihood (ML), the generalized method of moments (GMM), and the market-based (MB) method. The latter minimizes the spread or market volatilities obtained from the theoretical model and the data market, e.g., using a loss function (Bakshi, Cao, and Chen, 1997; and Bams, Lehnert, and Wolff, 2009). In this context, innovations in the computational science field offer considerable advantages to the financial industry, e.g., evolutionary computation, which has developed optimization methods based on evolutionary algorithms, representing an alternative to traditional optimization methods (either deterministic or stochastic) whenever the latter fail or the computational cost is high. Genetic algorithms are computational algorithms based on Darwinian evolution, which use operators such as reproduction, mutation, and selection to evolve a set of candidates or possible solutions to a given optimization problem. Gilli and Schumann (2011) provided a method using differential evolution to calibrate the Heston model and proposed only one loss function with artificially generated data. A notable finding is the utilization of alternative methods to decrease the risk model and the selection criteria of parameters. On the other hand, Cui, del Baño and Germano (2017) propose an efficient algorithm to calibrate the Heston model with nonlinear least squares using the Levenberg (1944) method by using the alternative representation of the characteristic function and avoiding discontinuities.

Soler and Kizys (2017) review some literature related to the use of evolutionary algorithms to solve classic and emerging problems in the field of finance. A non-exhaustive list of examples includes portfolio optimization, index tracking, enhanced indexation, credit risk, stock investments, financial project scheduling, option pricing, feature selection, bankruptcy and financial distress prediction, and credit risk assessment, they conclude that the use evolutionary algorithms in finance is becoming increasingly more popular among researchers from different communities in order to make efficient decisions. In Liu et al (2019) a method to calibrate stochastic volatility models using artificial neural networks is proposed that includes three phases: training, prediction and calibration. Some relevant results are the efficiency and precision of the calculations when compared with other methods such as differential
evolution that avoids getting stuck in a local minimum. In Palmer (2019) it is addressed a number of problems in computational finance, characterized by their high demand for computation time and resources. One of them is the calibration of the Heston model with variants of evolutionary algorithms such as particle swarm optimization (PSO) and differential evolution (DE). State-of-the-art DE algorithms were found to provide excellent calibration performance, and the prior use of rudimentary DE in the literature underestimates the use of these methods.

This paper is aimed at calibrating the parameters of Heston’s SVM with several loss functions by using two different methods: the first one minimizes a nonlinear function of several variables with constraints, and the second one employs the differential evolution algorithm. Both methods are applied to a set of implied volatilities on the Mexican Stock Exchange Index (IPC Spanish acronym for Indice de Precios y Cotizaciones) with four maturities and twenty-eight strike prices. The selection criterion for the parameters is minimizing the value of the mean square error of the implied volatility. It important to point out that the mean squared loss function weighted by the square of the vega reaches the minimum value by the first method with less error and time. Moreover, the empirical results show that for both methods the adjustment of implied volatilities is better for options with long-term maturities than those for short-term maturities, and the value of rho is negative, which means that the slope of the implied volatility curve is negative and is closely related to the asymmetry of the underlying asset returns distribution. The main conclusion is that a change in the slope of implied volatility reflects the supply and demand effects of options. Thus, high premium will move in the direction that market participants expect the underlying to move.

The structure of the paper is as follows: the next section provides a short review of SVMs; section 3 describes briefly the Heston model; section 4 describes the calibration of parameters using several loss functions; through section 5, the differential evolution algorithm to calibrate Heston’s model parameters is defined; in section 6, the Heston model parameters are calibrated by using loss functions according to the two methods of the previous sections, both methods are implemented for a set of implied volatilities of the IPC. Finally, in section 7, conclusions are given.

A Short Review of Stochastic volatility models

According to Garman’s (1976) theoretical proposal, an option price with stochastic volatility satisfy a second-order partial differential equation (PDE) with two fundamental variables: strike price and volatility. Volatility is not a trading asset, so arbitrage cannot eliminate volatility risk. Therefore, the market price is included in the PDE. Usually, the risk premium is zero or a constant volatility proportion. Final and boundary conditions for the option contracts state the premium. In these option pricing models;
underlying volatility is explicit and its parameters are obtained by statistical methods. The PDE solution from which the option price is obtained engender additional problems, since most of the time closed formulas are not available, and the least it can be asked for is an analytical solution as a prerequisite for the empirical work.

Wiggins’ (1987) work solves a PDE with a finite difference method for the logarithmic transformation of a state variable. Hull and White (1987) worked with a simple power series approximation technique based on the conditional underlying distribution of the average value of the stochastic variance. Stein and Stein (1991) and Heston (1993) proposed an analytic approach using the inverse Fourier transform. Of these three methods, the PDE numerical solution is the best known, but also the most computationally expensive. Hull and White’s (1987) and Stein and Stein’s (1991) approaches depend on the average stochastic variance distribution and assume instant correlation of state variables equal to zero, and Hull and White’s (1987) approach is more practical; however, Heston (1993) claimed that the Stein and Stein (1991) model works with the inverse Fourier transform even if the correlation is different from zero. On the other hand, Schöbel and Zhu (1999) developed a model with two correlated Brownian motions and estimated the characteristic function of the logarithmic price through a system of stochastic differential equations. Also, Zhu (2010) proposed a double root stochastic volatility model akin to Longstaff’s (1989) model. Date and Islyaev (2015) proposed an SVM based on a deterministic time structure modified by an escalated random variable. In this case, a closed approximation is obtained for a European-type option premium using higher order Greeks from its volatility.

**Revisiting Heston’s SVM**

An important feature in the Heston model is that the characteristic functions of the risk neutral probabilities are the solution to a second-order PDE (Venegas-Martínez, 2008). Using these probabilities, a formula is obtained to value a European call option under the assumption that volatility is stochastic. In this case, there is a correlation between the volatility and the underlying price, while the strike price is stated by the put-call parity. There are some complications with Heston’s model because it includes complex variables. Therefore, the option price is obtained by calculating the probability that the call option expires in the money. This probability is obtained by calculating the inverse of the characteristic function for the logarithm of the underlying price. The formalization of the Heston model is as follows. The underlying price $X_t$ is driven by the following stochastic differential equation:
\[
\frac{dX_t}{X_t} = \alpha \, dt + \sqrt{v_t} \, dB'_t
\]

where \( \alpha \) denotes the annual average return and \( B'_t \) is a Brownian motion or Wiener process (normal independent increments with zero mean and variance linearly depending of time). It is important noticing that Heston model uses the variance \( v_t \) instead of the volatility \( \sqrt{v_t} \). Indeed, the variance follows the Uhlenbeck and Ornstein (1930) process for volatility, denoted by \( g_t = \sqrt{v_t} \). Thus, \( dg_t = -\beta g_t \, dt + \delta dB'_2 \), where \( B'_2 \) is a Wiener process correlated to \( B'_t \), i.e., \( \text{Cov}(dB'_t, dB'_2) = \rho \). Using Itô’s lemma with \( v_t = g_t^2 \), the variance dynamics is obtained, and in this case, according to Cox, Ingersoll, and Ross (1985), the mean reversion process for the variance is:

\[
dv_t = \kappa (\theta - v_t) \, dt + \sigma \sqrt{v_t} \, dB'_2, \tag{1}
\]

where \( \kappa = 2\beta, \theta = \delta^2 / 2\beta, \) and \( \sigma = 2\delta \). If \( \kappa, \theta, \) and \( \sigma \) satisfy the conditions \( 2\kappa\theta > \sigma^2 \) and \( v_0 > 0 \). The above stochastic process is always positive, and the distribution of (1) is a non-central chi-squared. Moreover, the variance always has a mean reversion structure, where \( \kappa \) is the speed of adjustment toward the mean, \( \theta > 0 \), which is the long-term level of \( v_t \), and \( \sigma \) the instantaneous volatility of the variance. Intuitively, the parameter \( \theta \), in (1), is such that if the variance is become greater, then the variance is attracted toward \( \theta \). Regarding the estimation of the parameters, we have \( v_0 \) as an unobservable initial state variable (Bakshi et al., 1997). In this paper, \( v_0 \) is considered as a parameter to be estimated. To estimate the price of a European-type call option it is used:

\[
C_T (K) = e^{-rT} E \left[ (X_T - K, 0) \right] = e^{y_H} \Pi_1(y, v, \tau) - e^{-rT} K \Pi_2(y, v, \tau).
\tag{2}
\]

Here, \( \Pi_j(y, v, \tau) \) stands for the probability that the call option expires in the money (\( X_T > K \)). In this case, \( y_t = \ln S_t \) (natural log of the asset price), \( v_t \) is variance at time \( t \),
\( \tau = T - t \) is the term, and \( \Pi_j(y,v,\tau) = P(\ln X_T > \ln K), \quad j = 1,2. \) If the characteristic functions \( \varphi_1(\phi; y, v) \) and \( \varphi_2(\phi; y, v) \) related to \( \Pi_j \) are known, it is possible to obtain each \( \Pi_j \) from its characteristic function by using the inversion theorem (see Waller, Turnbull, and Hardin, 1995):

\[
\Pi_j = P(\ln X_T > \ln K) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \text{Re} \left[ \frac{e^{-i\phi \ln K}}{i\phi} \varphi_j(\phi; y, v) \right] d\phi
\]

(3)

At maturity, when \( t = T \), i.e., \( \tau = 0 \), the probabilities are constraints on the terminal condition \( \Pi_j(y, v, 0) = 1_{y > \ln K} \), which implies that at maturity, when \( X_T > K \), the probability that the call option is in the money is equal to 1. In summary, according to Heston (1993), the call option price is:

\[
C_T(K) = e^{ry} \Pi_1(y,v,\tau) - e^{-r\tau} K \Pi_2(y,v,\tau)
\]

(4)

where \( \Pi_1 \) and \( \Pi_2 \) are defined in (2), and the characteristic functions are given by:

\[
\varphi_j(\phi; y, v) = \exp \left( C_j(\tau, \phi) + D_j(\tau, \phi)v + i\phi x \right)
\]

(5)

\[
C_j = r\phi i + \frac{\kappa \theta}{\sigma^2} \left\{ \left( b_j - \rho \sigma \phi i + d_j \right) \tau - 2 \ln \left[ \frac{1 - g_J e^{d_j \tau}}{1 - g_J} \right] \right\}, \quad D_j = \frac{b_j - \rho \sigma \phi i + d_j}{\sigma^2} \left[ \frac{1 - e^{d_j \tau}}{1 - g_J e^{d_j \tau}} \right],
\]

\[
g_j = \frac{b_j - \rho \sigma \phi i + d_j}{b_j - \rho \sigma \phi i - d_j}, \quad d_j = \sqrt{\left( \rho \sigma \phi i - b_j \right)^2 - \sigma^2 (2 u J \phi i - \phi^2)}.
\]

where \( i = \sqrt{-1} \) is the imaginary unity, \( u_1 = 1/2, u_2 = -1/2, b_1 = \kappa + \lambda - \rho \sigma \), and \( b_2 = \kappa + \lambda \). Here, \( \lambda \) is the parameter that represents the volatility premium as a function of the asset price, time and volatility. To obtain the option price, it is necessary to estimate the integrals for \( \Pi_j \) in
(2), and this leads to a numerical integration method. To estimate the put option price, the put-call parity is used:

\[ V_T(K) = C_T(K) + Ke^{-rT} - X_T. \]

Another relevant property of this model is that implied volatilities are estimated by option prices generated according to the model, and these have volatility smile skew. Stock options regularly have a negative slope on the implied volatilities. It is remarkable that when calibrating the parameters of Heston’s model by using market data, the correlation parameter value is negative most of the time.

![Figure 1. Effect of the parameters on implied volatility with different ρ values. Source: own elaboration.](image)

The correlation parameter ρ determines the skew trend. If ρ > 0, there is a positive slope, otherwise there is a negative slope. This effect is shown in Figure 1 by generating implied volatilities using the following values: \(X = 100, \ r = 0.05, \ \tau = T - t = 0.25, \ \kappa = 2, \ \theta = 0.01, \ \lambda = 0, \ \nu_0 = 0.01,\) with a strike range oscillating from 90 to 110. If the volatility values increase the variance, the smile skew increases, i.e. a higher \(\sigma\) parameter increases the implied volatility curvature, this can be seen in Figure 2 for the same values but using \(\rho = 0\).
Figure 2. Effect of the parameters on implied volatility using different values of $\sigma$.

Source: own elaboration.

The parameters $\kappa$, $\theta$, and $\nu_0$ manage the smile level. The adjustment speed of the mean reversion, $\kappa$, also controls the curve up to some point, and using bigger values of $\kappa$, the implied volatility curve flattens, as seen in Figure 3.

Figure 3. Effect of the parameters on implied volatility using different values of $\kappa$.

Source: own elaboration.
When the values of $\theta$ and $v_0$ either increase or decrease, a parallel movement can be observed in the volatility smile, as seen in Figures 4 and 5. To graph the plots, the bisection algorithm is used with the BSM model to calculate the set of implied volatilities provided that prices are estimated using the Heston model.

Figure 4. Effect of the parameters on implied volatility with different values of $\theta$.

Source: own elaboration.

Figure 5. Effect of the parameters on implied volatility using different values of $v_0$.

Source: own elaboration.
Loss functions

One alternative method for calibrating the parameters of the Heston model is to use loss functions. This method minimizes the difference amongst market prices and estimated prices, or between the market-implied volatilities and the Heston model. The set of the estimated parameters \( \Theta \) corresponds to the values that minimize the value of the corresponding loss function in such a way that the model prices or the implied volatilities are close enough to the potential market prices. To do this, it is necessary to use a minimization algorithm with the following constraints: \( \kappa > 0, \theta > 0, \sigma > 0, v_0 > 0, \rho \in [-0.9, 0.9] \).

In this regard, the loss functions include trading asset prices at a point in time (or its implied volatility) leading to risk neutral parameters in the Heston model. Let us assume that there is a set of stock maturities \( M_t \) with \( \tau_t (t = 1, \ldots, T) \) and a set of strike prices \( K_k (k = 1, \ldots, M_K) \). For each combination maturity-strike \( (\tau_t, K_k) \), there is a market price \( P(\tau_t, K_k) = P_{tk} \) and the corresponding price obtained by Heston’s model \( P(\tau_t, K_k; \Theta) = P^\Theta_{tk} \). Moreover, it is possible to link each option to a weight \( \omega_{tk} \). There are several methods of defining a loss function, generally categorized as base prices and base-implied volatilities. The former minimizes the error between market quotes and the estimated parameters by a given model. The error is defined as the squared difference between market quotes and the model estimations, or the difference absolute value. It is also possible to use the relative error, i.e., estimating the parameters by using the loss function of the mean squared error (MSE) and minimizing:

\[
L_1 = \frac{1}{M} \sum_{t,k} \omega_{tk} \left( P_{tk} - P^\Theta_{tk} \right)^2 ,
\]

\[(7)\]

with respect to \( \Theta \), where \( M \) is the number of quotes. The other commonly used loss function is the so-called relative loss of the mean squared error (RLMSE) given by:

\[
L_2 = \frac{1}{M} \sum_{t,k} \omega_{tk} \left( \frac{P_{tk} - P^\Theta_{tk}}{P_{tk}} \right)^2 .
\]

\[(8)\]

Another way of defining the error is by taking the absolute value in (7) or (8).
A shortcoming of the loss function $L_1$ is that options with short maturity, out of the money, and with low value contribute little to the sum in (7); the optimization process adjusts options in the money to the long term, impacting other options. A solution to this issue is to use only options in the money in such way that in (7) call options are used for strike prices smaller than the underlying price and put options for strike prices greater than the underlying price. Another solution is the RLMSE in (8), but the problem with $L_2$ is the opposite. In fact, provided that $P_{tk}$ is in the denominator, options on the low-price market saturate the sum in (8), although this excess is repaired by allocating the weights $\omega_{tk}$ to the terms in the objective function. It turns out that the selection weight process is arbitrary.

The second category of loss functions is associated with minimizing the error between the implied volatilities and the parameters estimated by a theoretical model. This can be also specified as the relative difference, absolute difference or squared difference between the implied volatilities and those given by a theoretical model. This kind of loss function is sensitive because in the derivatives market, options quotes are made in terms of the implied volatility. The adjustment precision of the model is obtained by comparing the implied volatilities and the model volatility, i.e., the calibration parameters, using the loss function of the MSE of the implied volatility:

$$L_3 = \frac{1}{M} \sum_{t,k} \omega_{tk} \left( \hat{\sigma}_{tk} - \hat{\sigma}^\Theta_{tk} \right)^2,$$

where $\hat{\sigma}_{tk} = \hat{\sigma}(\tau_t, K_k)$ and $\hat{\sigma}^\Theta_{tk} = \hat{\sigma}(\tau_t, K_k; \Theta)$ are respectively the market-implied volatilities and the model volatilities. It is also possible to use the relative and absolute versions of $L_3$. The principal drawback of (9) is that it needs exhaustive numeric calculation, i.e., for each iteration, the optimization process estimates all the prices with the Heston model $P^\Theta_{tk}$, and then a root algorithm is used as a bisection algorithm to calculate the implied volatility $\hat{\sigma}^\Theta_{tk}$ in $P^\Theta_{tk}$ and thus finally is obtained

$$\left( \hat{\sigma}_{tk} - \hat{\sigma}^\Theta_{tk} \right)^2.$$ A solution to this problem is to use the loss function of the weighted MSE times squared vega (Christoffersen, Heston, and Jacobs, 2009), which is an approximation to the function in (9). In this case:
\[ L_4 = \frac{1}{M} \sum_{t,k} \left( \frac{P_{tk} - P_\theta}{v_{tk}^2} \right)^2, \]

where \( v_{tk}^2 \) is the option price sensitivity in the BSM model in relation to the market-implied volatility \( v_{tk} \) evaluated at maturity \( \tau_t \) at strike \( K_k \), defined as:

\[ v_{tk} = X e^{q \tau_t} \Phi(z_{tk}) \sqrt{\tau_t}, \quad \text{and} \quad z_{tk} = \frac{\ln \left( \frac{X}{K_k} \right) + \left( r - q + \frac{1}{2} v_{tk}^2 \right) \tau_t}{v_{tk} \sqrt{\tau_t}}, \]

Here, \( r \) is the continuously compounded interest rate, \( q \) is the continuously compounded dividend yield, and \( \Phi(\cdot) \) is the cumulative distribution function of a standard normal random variable.

The main improvement of the loss function \( L_4 \) is the time reduction in its execution, although there is a loss of precision in the calibration of parameters. Another advantage is that for each evaluation of the loss function, it is necessary to recalculate the option price. The analytic expression of the characteristic functions guarantees that these calculations are done efficiently.

**Differential evolution: an algorithm based on evolutionary computation**

While traditional optimization methods (i.e., mathematical programming) are limited tools for finding optimal solutions, heuristic algorithms, most of the time, offer a better alternative. In recent years, the blend of computer science and biology has given rise to a new paradigm for solving highly complex problems, the so-called evolutionary computation. There are three main paradigms that make up evolutionary computation: evolutionary strategies, evolutionary programing, and genetic algorithms. It is worth mentioning that genetic algorithms are the most common method of evolutionary computation used in finance Ponsich, López and Coello (2013).²

² An introduction to evolutionary computing is given in Simon (2013), in which several approaches to solving optimization problems are presented.
Genetic algorithms were developed by Holland (1975) and they use the process of natural selection to find the optimal solution through a series of populations that combine and mutate solutions to reach the optimum. Different optimization problems require different methods of mutation and combination of solutions (Bradshaw, Walshaw, Ierotheou, and Parrott, 2009, p. 28). Genetic algorithms are also known as evolutionary algorithms because they are inspired by the theory of evolution of Charles Darwin, and generally work with operators such as reproduction, crossover, and mutation.

The differential evolution (DE) paradigm was developed by Storn and Price (1995). Like many other optimization algorithms, the DE was motivated by the determination of coefficients using the polynomials of Tchebyshev\(^3\) and the optimization of the coefficients in the digital filter. The first publication of DE dates back to Storn and Price (1997). The DE algorithm is described in four distinct stages (Vollrath and Wendland, 2009): initial selection of the population, mutation, recombination, and selection. These stages are typical in most evolutionary algorithms; however, the DE is unique in the way the parameter sets are mutated. The stages of DE applied to calibration of Heston’s model are listed as follows:

a) Initial population. Generate an \(n \times M_p\) array of randomly chosen parameter values. In this case, \(n\) is the number of parameters and \(M_p\) is the number of members of the population. The rows of the matrix must satisfy the restrictions of the parameters. Lower and upper limits must be chosen in such a way that the random values for the parameters are feasible, v.g., \(-1 < \rho < 1, \sigma < 3\) and \(0 < \kappa < 10, \theta > 0, v_0 > 0\). For Step 2, it is necessary that \(M_p \geq 4\). This implies that with \(n = 5\):

\[
P = \begin{pmatrix} z_1 & z_2 & \cdots & z_{M_p} \end{pmatrix},
\]

(11)

and every member of \(P\) is a parameter vector \(\mathbf{z}_i = (\kappa_i, \theta_i, \sigma_i, v_0, \rho_i)^t\).

b) Mutation. For each member \(\mathbf{z}_i\) of the population \((i = 1, \ldots, M_p)\), randomly select another three members \(\mathbf{z}_{e1}, \mathbf{z}_{e2},\) and \(\mathbf{z}_{e3}\), different from each other and from \(\mathbf{z}_i\). This

---

\(^3\) A fast review of the polynomials of Tchebyshev can be seen in Venegas-Martínez (2008).
requires to select indexes $\varepsilon_1, \varepsilon_2$, and $\varepsilon_3$ such that $i \neq \varepsilon_1 \neq \varepsilon_2 \neq \varepsilon_3$. For every member, create a new giver member:

$$x_i = z_{\varepsilon_1} + F (z_{\varepsilon_2} - z_{\varepsilon_3}),$$

(12)

where $F$ is a constant mutation factor that controls the size of the mutations and takes values in $[0, 2]$. The giver members and components, $x_i$, are used to produce a candidate member that either replaces $z_i$ or not, depending on the outcomes of next steps.

c) Recombination. The candidate element $c_i$ is generated from the $x_{ij}$ elements of the giver and components of the member $z_{ij}$ ($j = 1, \ldots, 5$) according to the following: for each element $j$, a random number of a uniform variable $U_{ij}$ is generated and the candidate element is defined as follows:

$$c_{ij} = \begin{cases} x_{ij} & \text{if } U_{ij} \leq R, \text{ or } j = E \\ z_{ij} & \text{otherwise} \end{cases}$$

(13)

where $E$ is a randomly selected integer and $R$ is the denominated crossover ratio. $E$ ensures that the candidate is not the same as the member, i.e., $c_i \neq z_i$. This is because even if all $U_{ij}$ are small and the inequality $U_{ij} \leq R$ is not fulfilled at all, the corresponding element $x_{ij}$ will turn into the element $c_{ij}$ of the candidate $c_i$. This step is continued for each candidate parameter.

d) Candidate selection. The member $z_i$ and the candidate $c_i$ are both introduced into the objective function $f(z)$. If the candidate’s value is less than that of the member, the candidate replaces $z_i$ as the $i$th term member of the population $P$ in (11). Thus, the instruction is:
When all the elements of the population have admitted the procedure, the population is renovated, and the next generation starts over again in step 2. Therefore, the algorithm is executed for all generations, \( G \). In this way, the set of parameters is chosen for the member of the population in the final generation with the smallest value of \( f(z) \).

To further explain the above procedure, it is important to note that when evaluating a loss function given by equations (7)-(10) prices corresponding to a set of different strikes and maturities are calculated. In this case, the characteristic function \( \varphi \) only depends on the time to maturity, not on the strike price. Therefore, we can previously process the variables that are constant for a given maturity and then calculate the prices for all strike prices for this maturity. The following algorithm summarizes this.

**Algorithm 5.1 Computing option prices for a given surface.**

1. Set parameters, set \( \mathbb{C} = \) maturities, set \( \mathcal{K} = \) strikes
2. for \( \tau \in \mathbb{C} \) do
3. compute characteristic function \( \varphi \)
4. for \( X \in \mathcal{K} \) do
5. compute price for strike \( X \), maturity \( \tau \)
6. end for
7. end for
8. compute objective function \( L \)

Finally, the following algorithm summarizes the proposed methodology.

**Algorithm 5.2 Differential Evolution.**

1. set parameters \( M_p, N_G, F, R \)
2. initialize population \( P_{j,i}^{(z)}, j = 1, \ldots, p, i = 1, \ldots, M_p \)
\[ P = \left[ \kappa_{Lo} + (\kappa_U - \kappa_{Lo}) \times \text{rand}(1, M_P) \right] \times \ldots \]
\[ \theta_{Lo} + (\theta_U - \theta_{Lo}) \times \text{rand}(1, M_P) \times \ldots \]
\[ \sigma_{Lo} + (\sigma_U - \sigma_{Lo}) \times \text{rand}(1, M_P) \times \ldots \]
\[ v_{0Lo} + (v_{0U} - v_{0Lo}) \times \text{rand}(1, M_P) \times \ldots \]
\[ \rho_{0Lo} + (\rho_{0U} - \rho_{0Lo}) \times \text{rand}(1, M_P) \times \ldots \]

3. for k=1 to \( N_G \) do

4. \( P^{(x)} = P^{(z)} \)

5. for i=1 to \( M_P \) do

6. generate \( \varepsilon_1, \varepsilon_2, \varepsilon_3 \in \{1, \ldots, M_P\} \), \( \varepsilon_1 \neq \varepsilon_2 \neq \varepsilon_3 \neq i \)

7. Compute \( P^{(x)}_{ij} = P^{(z)}_{ij} + F \times \left( P^{(z)}_{i\varepsilon_2} - P^{(z)}_{i\varepsilon_3} \right) \)

8. for \( j = 1 \) to \( p \) do

9. if \( U < R \) then \( P^{(c)}_{j,i} = P^{(c)}_{j,i} \) else \( P^{(c)}_{j,i} = P^{(x)}_{j,i} \)

10. end for

11. if \( f(P^{(c)}_{ij}) < f(P^{(x)}_{ij}) \) then \( P^{(z)}_{ij} = P^{(c)}_{ij} \) else \( P^{(z)}_{ij} = P^{(x)}_{ij} \)

12. end for

13. end for

14. find member with lowest value \( h = \text{find}(f = \min, f(P^{(z)}_{ij})) \)

15. solution = \( P^{(z)}_{ih} \)

**Empirical Results**

This section considers a set of option prices for the Mexican Stock Exchange (BMV). The information was obtained from the Bulletin of the Summary of the Options Market since 06/26/2015, published by the Mexican Derivatives Market (MexDer). Four deadlines are considered: 84, 175, 266, and 357 days. The implied volatilities of the options traded were obtained from the price vendor, and the risk-free rate was acquired from the Mexican Central Bank (Banxico). Table 1 shows the relevant data for the calibration of the parameters.
Table 1
Relevant data for calibration

<table>
<thead>
<tr>
<th>Date</th>
<th>Expiration Date</th>
<th>( \tau )</th>
<th>Risk-free rate</th>
<th>IPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>26/06/2015</td>
<td>18/09/2015</td>
<td>0.2333</td>
<td>3.2950</td>
<td>45,566.33</td>
</tr>
<tr>
<td>26/06/2015</td>
<td>18/12/2015</td>
<td>0.4861</td>
<td>3.2950</td>
<td>45,566.33</td>
</tr>
<tr>
<td>26/06/2015</td>
<td>18/03/2016</td>
<td>0.7389</td>
<td>3.2950</td>
<td>45,566.33</td>
</tr>
<tr>
<td>26/06/2015</td>
<td>17/06/2016</td>
<td>0.9917</td>
<td>3.2950</td>
<td>45,566.33</td>
</tr>
</tbody>
</table>

Source: own elaboration with data from BMV, MexDer and Banxico.

Table 2 shows the expiration dates, strike prices (K), and implied volatilities of the IPC call options obtained from the price vector. There are 31 strike prices, of which 22 are options that are in the money and the remainder is out of the money. This is important since the loss functions previously presented give different results in the calibration under this assumption.

Table 2
Implied volatility of the IPC call options

<table>
<thead>
<tr>
<th>K</th>
<th>84/360</th>
<th>175/360</th>
<th>266/360</th>
<th>357/360</th>
</tr>
</thead>
<tbody>
<tr>
<td>35000</td>
<td>0.2465</td>
<td>0.2441</td>
<td>0.2324</td>
<td>0.2225</td>
</tr>
<tr>
<td>35500</td>
<td>0.2427</td>
<td>0.2386</td>
<td>0.2281</td>
<td>0.2192</td>
</tr>
<tr>
<td>36000</td>
<td>0.2395</td>
<td>0.2332</td>
<td>0.2239</td>
<td>0.2159</td>
</tr>
<tr>
<td>36500</td>
<td>0.2367</td>
<td>0.2280</td>
<td>0.2197</td>
<td>0.2126</td>
</tr>
<tr>
<td>37000</td>
<td>0.2340</td>
<td>0.2229</td>
<td>0.2154</td>
<td>0.2093</td>
</tr>
<tr>
<td>37500</td>
<td>0.2315</td>
<td>0.2178</td>
<td>0.2112</td>
<td>0.2060</td>
</tr>
<tr>
<td>38000</td>
<td>0.2290</td>
<td>0.2130</td>
<td>0.2070</td>
<td>0.2027</td>
</tr>
<tr>
<td>38500</td>
<td>0.2260</td>
<td>0.2081</td>
<td>0.2028</td>
<td>0.1994</td>
</tr>
<tr>
<td>39000</td>
<td>0.2225</td>
<td>0.2034</td>
<td>0.1985</td>
<td>0.1960</td>
</tr>
<tr>
<td>39500</td>
<td>0.2185</td>
<td>0.1987</td>
<td>0.1943</td>
<td>0.1926</td>
</tr>
<tr>
<td>40000</td>
<td>0.2138</td>
<td>0.1941</td>
<td>0.1901</td>
<td>0.1892</td>
</tr>
<tr>
<td>40500</td>
<td>0.2082</td>
<td>0.1897</td>
<td>0.1859</td>
<td>0.1857</td>
</tr>
<tr>
<td>41000</td>
<td>0.2019</td>
<td>0.1852</td>
<td>0.1817</td>
<td>0.1822</td>
</tr>
<tr>
<td>41500</td>
<td>0.1948</td>
<td>0.1808</td>
<td>0.1775</td>
<td>0.1786</td>
</tr>
<tr>
<td>42000</td>
<td>0.1871</td>
<td>0.1764</td>
<td>0.1734</td>
<td>0.1750</td>
</tr>
<tr>
<td>42500</td>
<td>0.1789</td>
<td>0.1720</td>
<td>0.1693</td>
<td>0.1713</td>
</tr>
<tr>
<td>43000</td>
<td>0.1703</td>
<td>0.1678</td>
<td>0.1653</td>
<td>0.1675</td>
</tr>
<tr>
<td>43500</td>
<td>0.1616</td>
<td>0.1636</td>
<td>0.1614</td>
<td>0.1637</td>
</tr>
<tr>
<td>44000</td>
<td>0.1530</td>
<td>0.1595</td>
<td>0.1576</td>
<td>0.1598</td>
</tr>
<tr>
<td>44500</td>
<td>0.1447</td>
<td>0.1554</td>
<td>0.1540</td>
<td>0.1559</td>
</tr>
</tbody>
</table>
45000  0.1370  0.1515  0.1506  0.1520
45500  0.1300  0.1478  0.1476  0.1481
46000  0.1239  0.1444  0.1451  0.1442
46500  0.1187  0.1415  0.1434  0.1406
47000  0.1145  0.1396  0.1428  0.1374
47500  0.1112  0.1399  0.1437  0.1351
48000  0.1088  0.1446  0.1469  0.1346
48500  0.1071  0.1541  0.1528  0.1365
49000  0.1060  0.1652  0.1599  0.1401
49500  0.1053  0.1758  0.1667  0.1444
50000  0.1051  0.1860  0.1731  0.1486

Source: own elaboration with data from BMV and MexDer.

Figure 6 shows the implied market volatility surface and the market prices of IPC options, computed from the information in the bulletin. Notice the reduction in the differences in prices, excluding options with little premium value; for the calibration the choice of the options is a settlement price greater than $100, i.e. prices from $K = 49000$ to $K = 50000$ are excluded. Therefore, there are 28 strike prices and 4 expiration dates, a total of 124 implied volatilities. An implied volatility bias is observed for each maturity level. Implied volatility is greater for small strike prices corresponding to call options in the money and put options out of the money, and implied volatility decreases as the strike price increases. Another feature is that this is more common for options with shorter maturities, and is flattened for options with greater expiration. The monotonic relationship between price and volatility shows that put options out of the money have a higher premium than the call option out of the money.

Figure 6. Surface of the implied market volatility and IPC market quotes.

Source: own elaboration with data from BMV, MexDer and Banxico.
Table 3 shows the parameters of the Heston model calibrated with several loss functions (MSE, RLMSE, and MSE of implicit volatility), the time in seconds that the calibration takes, and the value of each objective function. The initial values of the parameters are as follows: $\kappa = 9$, $\theta = 0.05$, $\sigma = 0.3$, $v_0 = 0.05$, $\rho = -0.8$. In this case, the minimum of a nonlinear function of several variables with constraints is determined, results of table 3 are obtained with the algorithm developed in Nelder and Mead (1965). The set of parameters that is chosen is the one that has the lowest value of MSE of implicit volatility, that is, by using the function $L_4$. In all three cases, it is observed that the value of $\rho$ is negative, which means that the slope of the implied volatility curve is negative, as it can be seen in Figure 7.

<table>
<thead>
<tr>
<th></th>
<th>$\kappa$</th>
<th>$\theta$</th>
<th>$\sigma$</th>
<th>$v(0)$</th>
<th>$\rho$</th>
<th>ECM-VI</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_1$</td>
<td>2.4240</td>
<td>0.0374</td>
<td>1.0590</td>
<td>0.0239</td>
<td>-0.3313</td>
<td>5.06E-05</td>
<td>32.66</td>
</tr>
<tr>
<td>$L_2$</td>
<td>15.0580</td>
<td>0.0274</td>
<td>1.9992</td>
<td>0.0140</td>
<td>-0.5583</td>
<td>7.67E-05</td>
<td>40.60</td>
</tr>
<tr>
<td>$L_3$</td>
<td>3.5429</td>
<td>0.0303</td>
<td>0.5518</td>
<td>0.0185</td>
<td>-0.7241</td>
<td>8.88E-05</td>
<td>4.574</td>
</tr>
<tr>
<td>$L_4$</td>
<td>2.6922</td>
<td>0.0351</td>
<td>0.8059</td>
<td>0.0197</td>
<td>-0.5215</td>
<td>3.65E-05</td>
<td>28.137</td>
</tr>
</tbody>
</table>

Source: own elaboration with data from BMV, MexDer and Banxico.

4 The non-linear problem is solved using the fmincon function in Matlab.
The above results show that the adjustment of implied volatilities is better for options with long-term maturities than for short-term maturities, and the time it takes to adjust is much lower compared to DE, as confirmed in Table 4.

Figure 8 shows the surface of implied market volatility and model volatility generated with the parameters obtained from $L_4$. In this case, 31 maturity terms are considered, varying from 84 to 357 days, and there are 271 strike prices, which range from every 50 points from K=35000 up to K=48500. The importance of the surface lies in the fact that while bonds are characterized by maturity, options are distinguished by their term at maturity and their strike price, for this reason requiring a surface instead of a curve. In addition, the yield curve on any given day is a concise description of bond market prices. Another usefulness of the surface is that it is feasible to value an exotic option with the information generated by the volatility surface; in the case that there are deadlines at maturity and strike prices that do not match the deadlines of the surface, an interpolation bilinear process is carried out.
Figure 8. Surface of implied market volatility and model volatility using the parameters obtained from function $L_4$.

Source: own elaboration with data from BMV, MexDer and Banxico.

Table 4 shows the parameters of the Heston model calibrated with the three previous loss functions but using the DE algorithm. Unlike Vollrath and Wendland (2009), we chose the number of population members equal to 45 times the number of parameters for calibration, in such a way that $M_p = 225$, the probability of the crossover ratio $R = 0.5$, the mutation factor $F = 0.8$, and the number of generations $G = 600$. The initial values of the parameters for the calibration are as follows: $\kappa = 9, \theta = 0.05, \sigma = 0.3, v_0 = 0.05, \rho = -0.8$. The set of parameters that is chosen is the one with the lowest value of the quadratic error of implied volatility which is the function $L_4$. In all three cases, it can be observed that the value is negative, which means that the slope of the implied volatility curve is negative, as it can be seen in Figure 9.

Table 4
Calibrated parameters with loss functions and differential evolution

<table>
<thead>
<tr>
<th>Function</th>
<th>$\kappa$</th>
<th>$\theta$</th>
<th>$\sigma$</th>
<th>$v(0)$</th>
<th>$\rho$</th>
<th>ECM-VI</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_1$</td>
<td>2.6498</td>
<td>0.0355</td>
<td>1.0073</td>
<td>0.0232</td>
<td>-0.3910</td>
<td>4.66E-05</td>
<td>210.6515</td>
</tr>
<tr>
<td>$L_2$</td>
<td>15.7152</td>
<td>0.0277</td>
<td>2.0000</td>
<td>0.0088</td>
<td>-0.5979</td>
<td>7.38E-05</td>
<td>203.7326</td>
</tr>
<tr>
<td>$L_3$</td>
<td>5.9857</td>
<td>0.0311</td>
<td>1.1976</td>
<td>0.0172</td>
<td>-0.5836</td>
<td>4.62E-05</td>
<td>28461.3763</td>
</tr>
<tr>
<td>$L_4$</td>
<td>4.6792</td>
<td>0.0322</td>
<td>1.0751</td>
<td>0.0185</td>
<td>-0.5487</td>
<td>4.15E-05</td>
<td>266.0566</td>
</tr>
</tbody>
</table>

Source: own elaboration with data from BMV, MexDer and Banxico.
Figure 9. Implied market and Heston volatilities estimated by function $L_4$ using differential evolution.

Source: own elaboration with data from BMV, MexDer and MexDer.

The above results indicate that the adjustment of implied volatilities is somewhat better for options with long-term maturities than for short-term maturities, and the time it takes to adjust is much greater in comparison with loss functions.\(^5\)

Finally, Figure 10 shows the adjustment for the area of implied market volatilities and the surface implied market volatility, generated with the parameters of $L_4$ using DE. As it can be seen, the implied volatilities for short maturities are more sensitive than those obtained by the first method, which greatly affects options prices within the money as their premiums are greater than options out of the money.

---

\(^5\) The non-linear problem is solved using the “fmincon” function in Matlab.
Conclusions

A relevant feature of the differential evolution algorithm is that it does not require initial parameters since it generates many tests of a population in parallel. Differential evolution performs sampling over the entire parameter space, and if it is carried out enough, it practically guarantees convergence to a global minimum. However, the disadvantage is that it requires a considerable number of repetitions, involving a high cost in time and computational resources.

Another possibility for calibrating the parameters of a stochastic volatility model is by using loss functions. This approach minimizes the distance that measures the error between market prices and model prices, or between implicit market volatilities and those form the theoretical model. In this framework, the Heston model generates consistent market volatility smiles for relatively distant maturities (six months and more). However, for short maturities, it presents problems as it does not generate consistent volatility smiles with the market; this depends to a large extent on the dynamics of the underlying asset, which affects the market quotations at a given time.

The effect of the variation of the parameters of the Heston model on the implied volatility curve has been examined in this research, and it is observed that the correlation parameter $\rho$ determines the direction of skew; $\rho > 0$ corresponds to a positive slope and $\rho < 0$ to a negative slope; while the volatility of the variance $\sigma$ determines the curvature of the smile. The parameters $\kappa$, $\theta$, and $v_0$ control the level of the smile.
The contribution of this study can be summarized as follows, the calibration of parameters of the Heston model with loss functions and differential evolution algorithm, in particular with a loss function of the weighted MSE times squared vega which employs option price sensitivity in the Black and Scholes model in relation to the market-implied volatility. An advantage of this procedure is that it is able to produce precise estimates, but at the expense of increased computation time.

In summary, in this investigation, the parameters of the Heston stochastic volatility model with loss functions have been calibrated using two different methods: the first determines the minimum of a nonlinear function of several variables with constraints; the second employs the algorithm of differential evolution. The two methods are applied to a set of IPC call options with four maturities and twenty-eight strike prices. The criterion for choosing the parameters is the one that has the minimum value of the MSE of implied volatility. The results for the two proposed methods indicate that the function of loss of the MSE weighted by the square of the vega is the one that has the minimum value, but with a minor error and time cost using the first method. In addition, in both methods the adjustment of implied volatilities is better for options with long-term maturities than for short-term maturities, and the $\rho$ value is negative, which means that the slope of the implied volatility curve is negative which reflects a skewed distribution of the underlying asset's return. This implies that the volatility modeled as a stochastic function of the underlying asset price is in line with assumption of stochastic volatility.

References


