Interdisciplinary Studies of Complex Systems
No. 12 (2018) 5–16

Volatility of prices of financial assets

Luca Di Persio,[1] *Nicola Gugole*[2]

1 Introduction

When it comes to analyze a financial time series, volatility modelling plays an important role. As an example, the variance of financial returns often displays a dependence on the second order moments and heavy-peaked and tailed distributions. In order to take into account for this phenomenon, known at least from the work of [22] and [14], econometric models of changing volatility have been introduced, such as the *Autoregressive Conditional Heteroskedasticity* (ARCH) model by Engle, see [13]. The idea behind the ARCH model is to make volatility dependent on the variability of past observations. Taylor, in [26], studied an alternative formulation in which volatility was driven by unobserved components, and has come to be known as the *Stochastic Volatility* (SV) model. Both the ARCH and the SV models, covered in Section 2, have been intensively studied in the past decades, together with more or less sophisticated estimation approaches, see [25], as well as concerning *concrete applications*, see, e.g., [9], and references therein.

Parallel to the study of discrete-time econometric models for financial time series, more precisely in the early 1970's, the world of option pricing experienced a great contribution given by the work of Fischer Black and Myron Scholes. The Black-Scholes (BS) model, see [4], assumes that the price of the underlying asset of an option contract follows a geometric Brownian motion. Latter type of approach has been also used within the framework of interest rate dynamics, see, e.g., [6], and references therein. One of the most successful extensions has been the continuous-time *Stochastic Volatility* (SV) model, introduced with the work of Hull and White, see, [19]. A major contribution was successively due to Heston in [18], indeed he developed a model which led to a quasi-closed form expression for European option prices. Differently from the BS model, the volatility is not longer considered constant, but it is allowed to vary trough time in a stochastic way. In Section 3 we will start from a sub-class of SV models, which is the one of *Local Volatility* (LV), being characterized by a deterministic time-varying volatility, and then we will consider the general SV case, providing information about the pricing equation as made, e.g., in [5] or, from a point of view more centred towards applications, in [12], and references therein.

[1] Department of Computer Science, University of Verona. luca.dipersio@univr.it
[2] Department of Computer Science, University of Verona.

2 Discrete-time models

Discrete-time models for the volatility, as said in the introduction, are born in order to analyze and reproduce the behavior of real financial time series, which are often characterized by a number of *stylized facts*, i.e., features of particular interest.

- The variance of returns of financial products is often subject to the so called *volatility clustering effect.* This means that the returns show an high serial autocorrelation: periods of high volatility are followed by periods with the same feature and viceversa.
- As noted in the pioneer works by Mandelbrot, see [22], and Fama, see [14], the variance of financial returns often displays a dependence on the second order moments and heavy-peaked and tailed distributions.
- Stock returns often exhibit the so called *leverage effect*: the conditional variance responds in an asymmetric way with respect to rises or falls of the asset price.
- The *covariation effect* captures the fact that the volatilities of different financial assets could be correlated: large changes in the returns of an asset can induce a similar behavior in other assets.

In the following we will briefly introduce the ARCH model, see [13], trying to emphasize its limits. Then, we will treat the SV model, see [26], and related extensions, in order to model the aforementioned stylized facts. It is worth to mention that different, more numerically oriented methods, can be also fruitfully exploited, as, e.g., suggested in [10, 11] and references therein.

2.1 ARCH model

One of the most popular discrete-time models for the stochastic volatility is the ARCH model, which establishes a connection between the past squared returns of a financial asset and its current conditional variance. We let $\{y_t\}_{t=1}^{\infty}$ be the return process of some observation model. In the original formulation of Engle, see [13], the dynamic of the ARCH(1) was given by

$$y_t|\mathcal{F}_{t-1} \sim N(\mu, \sigma_t^2), \tag{1}$$

$$\sigma_t^2 = w + \alpha y_{t-1}^2, \tag{2}$$

where $w, \alpha \geq 0$ are real non-stochastic parameters, $\mathcal{F}_t$ denotes the global information up to time t. Naturally, eq. (2) could be generalized to the general ARCH(p) case

$$\sigma_t^2 = w + \sum_{i=1}^{p} \alpha_i y_{i-1}^2, \quad \alpha_i \geq 0,$$

in which the conditional variance is given by a linear combination of p-lagged squared error terms. As noted by Nelson, see [23], the ARCH model presents at least 2 drawbacks:

- Constraints must be imposed on the parameters in order to guarantee the positivity of the conditional variance, however they are often violated in the classical estimation procedures.

- It is not possible to model the conditional variance as a random oscillatory process, which is a recurrent situation observed in real data.

In the following we will present the *Stochastic Volatility* (SV) model due to Taylor, see [26] and [27], and able to overcome the aforementioned difficulties.

2.2 Stochastic volatility (SV) model

The peculiarity of the SV model by Taylor is that the variance of the returns is modeled as an unobserved process. In [27] Taylor shows that this model can be transposed into a continuous time version, useful when it comes to price options and other modern financial instruments.

Denoting again $\{y_t\}_{t=1}^{\infty}$ as the return process of some observation model, the SV parametrization sets

$$\begin{cases} y_t = \exp(h_t/2)\varepsilon_t, & \varepsilon_t \sim N(0,1), \\ h_t = w + \alpha h_{t-1} + \eta_t, & \eta_t \sim N(0,\sigma_\eta^2), \end{cases} \tag{3}$$

where the ε_t's and the η_t's are independent. Notice that $\{h_t\}_{t=1}^{\infty}$ represents nothing but the logarithm of the volatility of the return process $\{y_t\}_{t=1}^{\infty}$. In this way, the positivity of the related variance is guaranteed. α can be seen as a persistence parameter. Notice that $\{h_t\}_{t=1}^{\infty}$ is a standard autoregressive AR(1) process only when $|\alpha| < 1$, case in which it is strictly stationary with mean an variance

$$\mu_h = \mathbb{E}[h_t] = \frac{w}{1-\alpha}, \quad \sigma_h^2 = \mathrm{Var}(h_t) = \frac{\sigma_\eta^2}{1-\alpha^2}\,.$$

Equation (3) is not the unique way to write the dynamic of the model, see [24] for equivalent formulations. In particular, the SV model can be extended in order to take into account the following stylized facts, see [21] for further details:

- In some cases, the kurtosis of a financial time series is greater than 3. This corresponds to fatter tails with respect to a normal distribution. The problem can be solved by allowing ε_t in equation (3) to have a Student t-distribution.
- A financial asset can exhibit the so called *leverage effect*, that is, the volatility responds in an asymmetric way to rises or falls in the returns. This fact can be incorporated in the SV model by introducing a negative instantaneous correlation between ε_t and η_t in equation (3).

2.2.1 Estimation procedures

Differently from the ARCH-type models, we do not know the conditional distribution of y_t in closed form, see equation (1). For this reason, the standard Maximum Likelihood (ML) approach is hard to implement. Indeed, if we denote by $\mathbf{y} = (y_1, \ldots, y_N)$ the vector of N consecutive observations of the process y_t, by $\mathbf{h} = (h_1, \ldots, h_N)$ the corresponding vector for the log-volatilities, and by $\theta = (w, \alpha, \sigma_\eta^2)$ the vector of parameters, then the likelihood can be written as

$$\mathcal{L}(\mathbf{y};\theta) = \int p(\mathbf{y},\mathbf{h}|\theta)\,d\mathbf{h} = \int p(\mathbf{y}|\mathbf{h},\theta)p(\mathbf{h}|\theta)\,d\mathbf{h}, \tag{4}$$

where we integrate with respect to the joint probability distribution of the data. The N-dimensional integral in equation (4) requires the use of computationally involved numerical methods and for this reason the estimation of the parameters is hard. Following [24], we briefly cite some alternative estimation procedures:

- *Generalized Method of Moments* (GMM): this method was introduced by Taylor, see [26]. The basic idea is to match the empirical moments of the observed vector $\mathbf{y}$ with the corresponding theoretical ones, which can be computed explicitly, hence the key advantage is that the conditional distribution of y_t is not required. More precisely, we need to minimize the objective function $\mathbf{Q} = \mathbf{g}'\mathbf{W}\mathbf{g}$ with respect to the vector of parameters θ, where

$$\mathbf{g} = \frac{1}{N}\left(\sum_{i=1}^{N} y_i^2 - \mathbb{E}[y_i^2], \sum_{i=1}^{N} y_i^4 - \mathbb{E}[y_i^4], \sum_{i=2}^{N} y_i^2 y_{i-1}^2 - \mathbb{E}[y_i^2 y_{i-1}^2], \right.$$
$$\left. \ldots, \sum_{i=\tau+1}^{N} y_i^2 y_{i-\tau}^2 - \mathbb{E}[y_i^2 y_{i-\tau}^2]\right)^\top, \quad \tau \geq 1,$$

  and $\mathbf{W}$ is a positive definite, symmetric weighting matrix of dimension $(\tau+2)\times(\tau+2)$. It is possible to minimize $\mathbf{Q}$ using standard numerical routines.
- *Quasi-Maximum Likelihood estimation* (QML): this approach is based on the linearization of the SV model in equation (3). Assuming $\varepsilon_t \sim N(0,1)$ and defining $w_t = \log y_t^2$, it is possible to prove that

$$\begin{cases} w_t = -1.2704 + h_t + \xi_t, \\ h_t = w + \alpha h_{t-1} + \eta_t, \qquad \eta_t \sim N(0, \sigma_\eta^2), \end{cases} \tag{5}$$

  where $\xi_t = \log \varepsilon_t^2 - \mathbb{E}[\log \varepsilon_t^2]$, $\mathrm{Var}(\xi_t) = \pi^2/2$. Even if the errors ξ_t do not have a normal distribution, the underlying idea of the QML approach is to suppose $\xi_t \sim N(0, \pi^2/2)$ i.i.d., and to apply the Kalman filter to equation (5) in order to produce one-step ahead forecasts of w_t as well as h_t. Decomposing the prediction error, it is possible to construct the Gaussian likelihood of the data, to be minimized in order to estimate the vector of parameters θ, see [17].

2.2.2 The multivariate case

A stylized fact which can not be captured by the standard univariate SV model is the so called *covariation effect*, that is, roughly speaking, the presence of a correlation between the volatilities of different financial series. Often, large changes in the returns of an asset are followed by large changes in other ones. This can be due to the presence of common unobserved factors influencing the dynamics of a set of assets. Volatilities are also subject to the coming of new information, such as trading volume, quote arrivals, goverment's health, dividend announcements and so on. All these phenomena suggest that a multivariate model could be better than an univariate one in term of adherence to real data.

The first multivariate SV model was proposed in [16]. We denote by $\mathbf{y}_t = (y_{1,t}, \ldots, y_{N,t})^\top$ the vector of returns related to N different assets at time t. The dynamic of the i-th component is assumed to be

$$\begin{cases} y_{i,t} = \exp(h_{i,t}/2)\varepsilon_{i,t}, \\ h_{i,t} = w_i + \alpha_i h_{i,t-1} + \eta_{i,t}, \end{cases}$$

where $\varepsilon_t = (\varepsilon_{1,t}, \ldots, \varepsilon_{N,t})$ and $\eta_t = (\eta_{1,t}, \ldots, \eta_{N,t})$ are mutually independent and normally distributed. Moreover

$$\mathrm{Var}(\eta_t) = \Sigma_\eta,$$

$$\mathrm{Var}(\varepsilon_t) = \Sigma_\varepsilon = \begin{pmatrix} 1 & \rho_{1,2} & \ldots & \rho_{1,N} \\ \rho_{1,2} & 1 & \ldots & \rho_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1,N} & \rho_{2,N} & \ldots & 1 \end{pmatrix},$$

where $|\rho_{i,j}| < 1$, so that Σ_ε is a correlation matrix. The weakness of the model is that it does not allow the covariances of the assets to evolve in an independent manner of the variances. If $i \neq j$,

$$\mathrm{Cov}(y_{i,t}, y_{j,t}|\mathbf{h}_t) = \mathbb{E}[y_{i,t}^2 y_{j,t}^2|\mathbf{h}_t] = \rho_{i,j} \exp\left(\frac{h_{i,t}}{2}\right) \exp\left(\frac{h_{j,t}}{2}\right),$$

and since

$$\mathrm{Var}(y_{i,t}|\mathbf{h}_t) = \exp\left(h_{i,t}\right),$$

it follows that the model has constant correlations, which can be a limiting fact in some situations, see, e.g., [25]. As in the univariate case, it is possible to estimate the parameters through a QML approach, see [16], by linearizing the corresponding equations.

The multivariate SV model admits also other representations, e.g., the factorial one, see [20]. The main advantage with respect to the previous multivariate model, is the reduction of the dimensionality of the parameter space: the returns vector $\mathbf{y}_t = (y_{1,t}, \ldots, y_{N,t})^\top$ is a linear combination of unobserved and common factors following a univariate SV dynamic. If we denote by $\mathbf{f}_t = (f_{1,t}, \ldots, f_{K,t})^\top$ the set of common factors at time t, then

$$\mathbf{y}_t = B\,\mathbf{f}_t + \mathbf{w}_t\,,$$

$$\begin{cases} f_{i,t} = \exp(h_{i,t}/2)\varepsilon_{i,t}, \\ h_{i,t} = \mu_i + \phi_i h_{i,t-1} + \eta_{i,t}, \end{cases} \qquad i = 1, \ldots, K,$$

where B is a constant matrix of dimension $N \times K$, $K < N$, $\mathbf{w}_t \sim N(0, \Omega)$ is the error vector and it is assumed independent of all the other term. The random variables $\varepsilon_{i,t}$ and $\eta_{i,t}$ are serially and mutually independent and normally distributed. We assume also that $|\phi_i| < 1$ so that the factor log-volatility processes $h_{i,t}$ are stationary. For more details about the model, see [20], [24].

3 Continuous-time models

In the early 1970's the world of option pricing experienced a great contribution given by the work of Fischer Black and Myron Scholes. They developed a new mathematical model to treat certain financial quantities publishing the related results in the article *The Pricing of Options and Corporate Liabilities*, see [4]. The latter work became soon a reference point in the financial scenario. Nowadays, many traders still use the Black and Scholes (BS) model to price as well as to hedge various types of contingent claims. An important property of the BS model is that all the involved parameters are not influenced by the risk preferences of investors. In particular, the BS approach is based on the so-called risk-neutral pricing assumption which greatly simplifies the associated derivatives analysis.

In particular, in the classical BS-model, the volatility parameter, let us indicate it with σ, is assumed to be constant. Latter hypothesis cannot be considered realistic, as simple empirical analyses can easily show. In particular it is rather simple to show that the implied volatility of a financial asset is not constant but varies with time to maturity $T > 0$, and with respect to the strike price K. Such a fact has started to become more and more evident since the general market crash in 1987. As a consequence, the real values of the volatility parameter that can be observed in the market do not give rise to a flat shape as the BS-model forecasts. In fact, if we fix the strike price value and we look at the corresponding implied volatility section, e.g., with respect to a plain vanilla option, the typical figure that appears justifies the definition of the so-called *smile/smirk effect.* The latter because, especially for short maturities, the implied volatility sections assume a shape which resembles a *smile* or a *smirk.*

As a consequence of the BS-model lack of description accuracy, new models have been developed to overcome issues of the type mentioned so far. This has been also produced approaches able to treat the increasingly complexity characterizing modern financial instruments. Between such alternatives to the BS analysis, we focus our attention on the so called *local volatility* (LV) and *stochastic volatility* (SV) models.

3.1 Local volatility models

The LV models can be seen as the simplest extension of the classical BS case, in order to achieve an exact reproduction of the volatility smile, through calibration to market data. The main difference is that in LV models, the instantaneous volatility is, in general, a function of the current time and the current asset price. If we denote by S_t the price of the asset at time t, we can write the related SDE as

$$dS_t = \mu(t, S_t)S_t\, dt + \sigma(t, S_t)S_t\, dW_t \ ,$$

where $S_0 > 0$, $\mu(t, S_t)$ is the instantneous drift, $\sigma(t, S_t)$ is the instantaneous volatility at time t, and W_t a Brownian motion. If $\sigma(t, S_t) \equiv \sigma > 0$ then we turn back to the BS case.

The first LV model appeared in the literature is the so called *Constant Elasticity of Variance* (CEV) model, see [7]. The latter is characterized by a volatility defined as

$$\sigma(t, S_t) \doteq \sigma S_t^{\gamma-1}, \quad \sigma > 0,$$

where γ must be determined with a calibration to market data. With $\gamma = 1$ we find the BS model, while $\gamma = 0$ leads to normally distributed returns.

3.1.1 The pricing equation

Denoting by $C = C(t, S_t; T, K)$ the time-t price of a vanilla option having as underlying the asset price S_t, maturity T and strike $K > 0$, then it is possible to show, assuming existence and uniqueness of the risk-neutral measure, that C solves the following PDE:

$$\frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + \frac{1}{2}\sigma(t, S)^2 S^2 \frac{\partial^2 C}{\partial S^2} = rC \ , \tag{6}$$

where $r > 0$ is the constant instantaneous spot rate, to be coupled with appropriate boundary conditions, depending on the nature of the option of interest. In particular, setting $C(T, S_T)$ to the option's payoff and solving the equation backwards from T to t, it is possible to find $C(t, S_t)$.

3.1.2 The Dupire formula

Suppose to have a set of vanilla option's prices related to time t. Is there a way to set $\sigma(t, S)$ in such a way to perfectly fit these prices? The answer is yes, and comes from the well known *Dupire formula*, see [3], [15], or [8]:

$$\sigma(T, K)^2 \doteq \sigma(T, K; t, S_t)^2 = 2\ \frac{\dfrac{dC}{dT} + rK\dfrac{dC}{dK}}{K^2\dfrac{d^2C}{dK^2}}\ . \tag{7}$$

In particular, if equation (7) holds at time $t = 0$, then the model is automatically calibrated to the initial market volatility smile. Moreover, it is possible to show that the right hand side of equation (7) is always well defined if the real market is arbitrage free. Manipulating a little bit the Dupire formula, we can rewrite it in the following way:

$$\frac{\partial C}{\partial T} + rK\frac{\partial C}{\partial K} - \frac{1}{2}\sigma(T, K)^2 K^2 \frac{\partial^2 K}{\partial K^2} = 0\ . \tag{8}$$

Equation (8) is similar to (6) in many aspects, however must be solved forward in order to find option's prices for all the values of K and T, fixing t and S_t.

Suppose, for simplicity of exposition, that $r = 0$. Then the Dupire formula (7) turns into

$$\sigma(T, K)^2 = 2\ \frac{\dfrac{\partial C}{\partial T}}{K^2\dfrac{\partial^2 C}{\partial K^2}}\ . \tag{9}$$

Usually, vanilla option prices are quoted in terms of the BS implied volatility $\sigma_{BS} = \sigma_{BS}(t, S_t; T, K)$, i.e., that value of the volatility which, once inserted into the BS pricing formula, gives the market price:

$$C(t, S_t; K, T) = C_{BS}(t, S_t; K, T, \sigma_{BS}).$$

By using the chain differentiation rules and the formulas of the BS greeks, it is possible to write equation in terms of σ_{BS}, instead if C, see [15], i.e.,

$$\sigma(\tau, K)^2 = \frac{2\frac{\partial \sigma_{BS}}{\partial \tau} + \frac{\sigma_{BS}}{\tau}}{K^2\left(\frac{\partial^2 \sigma_{BS}}{\partial K^2} - d_1\sqrt{\tau}\left(\frac{\partial \sigma_{BS}}{\partial K}\right)^2 + \frac{1}{\sigma_{BS}}\left(\frac{1}{K\sqrt{\tau}} + d_1\frac{\partial \sigma_{BS}}{\partial K}\right)^2\right)},$$

where $\tau \doteq T - t$ and

$$d_1 \doteq \frac{1}{\sigma_{BS}\sqrt{\tau}} \ln\left(\frac{S_t}{K}\right) + \frac{1}{2}\sigma_{BS}\sqrt{\tau}\ .$$

As a particular case, suppose that σ_{BS} is independent of K, i.e., the volatility smile has no skew, so $\sigma(\tau, K) = \sigma(\tau)$, where

$$\sigma(\tau)^2 = 2\tau\sigma_{BS}\frac{\partial \sigma_{BS}}{\partial \tau} + \sigma_{BS}^2 \ = \frac{\partial}{\partial \tau}\left(\tau\sigma_{BS}^2\right),$$

from which

$$\int_0^\tau \sigma(u)^2 du = \tau\sigma_{BS}^2\ .$$

3.2 Stochastic volatility models

The SV models represent a natural extension of the LV models. We will consider the following couple of SDEs:

$$\begin{cases} dS_t = \mu(t, S_t)S_t\, dt + \sqrt{v_t}S_t\, dW_t \\ dv_t = \alpha(t, S_t, v_t)\, dt + \eta\beta(t, S_t, v_t)\sqrt{v_t}\, dZ_t\,, \end{cases} \qquad \mathbb{E}[dW_t dZ_t] = \rho\, dt\,, \tag{10}$$

where η is the volatility of volatility, ρ represents the instantaneous correlation between the two Brownian motions W_t and Z_t, and $\gamma > 0$. In the limit $\eta \to 0$, we retrieve the SV case.

The Heston model is, nowadays, the most know SV model; it was introduced for the first time in [18]. Starting from equation (10), the Heston model corresponds to the choice

$$\alpha(t, S_t, v_t) \doteq \theta(\bar{v} - v_t), \quad \bar{v} > 0, \quad \theta > 0,$$

$$\beta(t, S_t, v_t) \equiv 1.$$

In other words, v_t is a Cox-Ingersoll-Ross (CIR) process, where $\bar{v}$ is the so called *long term mean* and θ represents the *speed of reversion.* This terminology reflects the fact that, for sufficiently large times, v_t will move around the value $\bar{v}$ with an intensity depending on the magnitude of $\bar{v}$. An important feature of the CIR process is, under some conditions on parameters, the positivity: in particular, we have to impose $2\theta\bar{v} \geq \eta^2$.

3.2.1 The pricing equation

In the BS case, as well as in the SV case, there is only one source of randomness, more precisely the process W_t, but in the SV case we have also random changes in the volatility to be hedged. The idea is to set up a portfolio containing the option of interest, a quantity $-\Delta_1$ of the underlying asset and a quantity $-\Delta_2$ of another asset depending on the volatility value v_t. Differentiating the portfolio value and imposing the usual risk-free conditions (random terms equal to zero and return equal to r), see [15] for further details, we end up with the following PDE:

$$\begin{aligned} &\frac{\partial C}{\partial t} + \frac{1}{2} v_t S_t^2 \frac{\partial^2 C}{\partial S^2} + r S_t \frac{\partial C}{\partial S} + \rho \eta v_t \beta S_t \frac{\partial^2 C}{\partial v \partial S} + \frac{1}{2} \eta^2 v_t \beta^2 \frac{\partial^2 C}{\partial v^2} \\ &= rC - (\alpha - \phi \beta \sqrt{v_t}) \frac{\partial C}{\partial v} \ , \end{aligned} \tag{11}$$

where $\phi = \phi(t, S_t, v_t)$ is the so called *market price of volatility risk*, and can be seen as the extra return (required by the investors) per unit of volatility risk. Defining

$$\tilde{\alpha} \doteq \alpha - \phi \beta \sqrt{v_t}$$

as the drift of the volatility v_t process under the risk-neutral measure, we could rewrite equation (11) in a more compact way as

$$\frac{\partial C}{\partial t} + \frac{1}{2} v_t S_t^2 \frac{\partial^2 C}{\partial S^2} + r S_t \frac{\partial C}{\partial S} + \rho \eta v_t \beta S_t \frac{\partial^2 C}{\partial v \partial S} + \frac{1}{2} \eta^2 v_t \beta^2 \frac{\partial^2 C}{\partial v^2} = rC - \tilde{\alpha} \frac{\partial C}{\partial v} \ . \tag{12}$$

Equation (12) is a good point to start with, if the aim is to calibrate the SV model to option prices, which are closely connected to the risk-neutral measure. In particular, we can assume that the SV model of interest, once fitted the related parameters to option prices, generates the risk-neutral measure such that the market price of volatility risk ϕ is equal to zero. This assumption makes sense when we are interested only in the pricing part, not in the statistical properties, which are described by the physical measure.

3.2.2 Calibrating the parameters of the Heston model

The main advantage of the Heston model with respect to other (potentially more realistic) stochastic volatility models is the existence of a fast and easily implemented quasi-closed form solution for European options, see [15] for the derivation. This computational efficiency in the valuation of European options becomes useful when calibrating the model to real option prices. How can we perform the calibration? The simplest way is to minimize the distance between the observed European call option prices and the theoretical ones. If we denote by θ the set of parameters of the Heston model, then we have to solve the non-linear least squares

$$\bar{\theta} = \arg\min_{\theta \in \Theta} \sum_{i=1}^{N} \Big(C_i^{obs} - C_i(\theta) \Big)^2 , \tag{13}$$

where $C_i^{obs} = C^{obs}(K_i, T_i)$, $i = 1, \ldots, N$, is the set of observed option prices, while $C_i(\theta) = C_i(K_i, T_i; \theta)$, $i = 1, \ldots, N$, is the set of option prices produced by the Heston model, and Θ denotes the parameter space. Alternatively, one could perform the minimization in equation (13) using a dataset of implied volatilities instead of the corresponding quoted option prices.

A different approach is adopted, for instance, in [1], and it is based on the Maximum Likelihood method. We can imagine the stock price S_t at time t as a function of a vector of state variables X_t following a multivariate stochastic volatility dynamic as in equation (10), i.e., $S_t = f(X_t)$ for some function f. Usually, either the stock price itself (or its logarithm) is taken as one of the state variables, hence we write $X_t = (S_t, Y_t)^\top$, where Y_t is the remaining set of state variables of length N. In general, part of the state vector X_t can not be directly observed. In [1], the idea is to assume that both a time series of stock prices and a vector of quoted option prices are observed. The latter vector at time t is denoted by C_t, and must be used in order to infer the time series for Y_t. If Y_t is multidimensional then a sufficient number of different option prices is needed. Roughly speaking, there are two ways to extract the value of Y_t from observed data:

- The first method is to compute option prices as a function of S_t and Y_t, for each parameter vector considered during the estimation procedure. In this way it is possible to identify the parameters both under the physical measure and the risk-neutral one.
- The second method consists in using the BS implied volatility as a proxy for the instantaneous volatility of the stock. This is a simplifying procedure, and it can be applied only in the case of a single stochastic volatility state variable.

Since, in general, the transition likelihood function for a stochastic volatility model is not known in closed form, then an approximation method must be used, see [2]. In this way it is possible to express, in an approximate closed form, the joint likelihood of X_t. Then, in order to find the likelihood of $(S_t, C_t)^\top$, which is entirely observed, it is necessary to multiply the likelihood of the vector X_t by an appropriate jacobian term. The last step is not necessary when a proxy for Y_t is used. For further details, see [1].

References

[1] Aıt-Sahalia, Yacine and Robert Kimmel. 2007. Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics* 83(2):413–452.

[2] Aıt-Sahalia, Yacine et al. 2008. Closed-form likelihood expansions for multivariate diffusions. *The Annals of Statistics*, 36(2):906–937.

[3] Bergomi, L. 2015. *Stochastic Volatility Modeling.* CRC Press.

[4] Black, F. and M. Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81(3):637–654.

[5] Cordoni, F. and L. Di Persio. 2014. First order correction for the characteristic function of a multidimensional and multiscale stochastic volatility model. *International Journal of Pure and Applied Mathematics* 93(5):741–752.

[6] Cordoni, F. and L. Di Persio. 2015. Invariant measure for the Vasicek interest rate model in the Heath–Jarrow–Morton–Musiela framework. *Infinite Dimensional Analysis, Quantum Probability and Related Topics* 18(3):1550022.

[7] Cox, J. C. and S. A. Ross. 1976. The valuation of options for alternative stochastic processes. *Journal of Financial Economics* 3(1–2):145–166.

[8] Derman, E. and M. B. Miller. 2016. *The Volatility Smile.* John Wiley & Sons.

[9] Di Persio, L., A. Cecchin, and F. Cordoni. 2017. Novel approaches to the energy load unbalance forecasting in the Italian electricity market. *Journal of Mathematics in Industry* 7(5).

[10] Di Persio, L. and M. Frigo. 2015. Maximum likelihood approach to Markov switching models. *WSEAS Transactions on Business and Economics* 12:239–242.

[11] Di Persio, L. and M. Frigo. 2016. Gibbs sampling approach to regime switching analysis of financial time series. *Journal of Computational and Applied Mathematics* 300:43–55.

[12] Di Persio, L. and S. Vettori. 2014. Markov switching model analysis of implied volatility for market indexes with applications to S&P 500 and dax. *Journal of Mathematics.*

[13] Engle, Robert F. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society* 50(4):987–1007.

[14] Fama, E. 1965. The behavior of stock market prices. *Journal of Business* 38:34–105.

[15] Gatheral, J. 2011. *The volatility surface: a practitioner's guide*, vol. 357. John Wiley & Sons.

[16] Harvey, Andrew, Esther Ruiz, and Neil Shephard. 1994. Multivariate stochastic variance models. *The Review of Economic Studies* 61(2):247–264.

[17] Harvey, Andrew C. 1990. *Forecasting, structural time series models and the Kalman filter.* Cambridge University Press.

[18] Heston, S. L. 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of financial studies* 6(2):327–343.

[19] Hull, John and Alan White. 1987. The pricing of options on assets with stochastic volatilities. *The Journal of Finance* 42(2):281–300.

[20] Jacquier, Eric, Nicholas G. Polson, and Peter E. Rossi. 1999. *Stochastic volatility: Univariate and multivariate extensions.*

[21] Jacquier, Eric, Nicholas G. Polson, and Peter E Rossi. 2004. Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics*, 122(1):185–212.

[22] Mandelbrot, Benoit. 1963. The variation of certain speculative prices. *The Journal of Business* 36(4):394–419.

[23] Nelson, Daniel B. 1991. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society* 59(2):347–370.

[24] Plataniotí, K., E. J. McCoy, and D. A. Stephens. 2005. A review of stochastic volatility: univariate and multivariate models. Technical report, working paper.

[25] Shepard, N. 1995. Statistical aspects of arch and stochastic volatility.

[26] Taylor, Stephen J. 1986. *Modelling financial time series.* World Scientific.

[27] Taylor, Stephen J. 1994. Modeling stochastic volatility: A review and comparative study. *Mathematical Finance* 4(2):183–204.