Stochastic Volatility and Option Pricing with Long-Memory in Discrete and Continuous Time

Alexandra Chronopoulou[1] Frederi G. Viens[2]

[1] Institut Élie Cartan, INRIA Nancy Grand-Est
615, rue du Jardin Botanique, 54600 Villers-lés-Nancy, France.
Alexandra.Chronopoulou@iecn.u-nancy.fr

[2]Department of Statistics, Purdue University,
150 N. University St., West Lafayette, IN 47907-2067, USA.
viens@purdue.edu

July 16, 2011

Abstract

It is commonly accepted that certain financial data exhibit long-range dependence. We consider a continuous time stochastic volatility model in which the stock price is Geometric Brownian Motion with volatility described by a fractional Ornstein-Uhlenbeck process. We also study two discrete time models: a discretization of the continuous model via an Euler scheme and a discrete model in which the returns are a zero mean iid sequence where the volatility is a fractional ARIMA process. We implement a particle filtering algorithm to estimate the empirical distribution of the unobserved volatility, which we then use in the construction of a multinomial recombining tree for option pricing. We also discuss appropriate parameter estimation techniques for each model. For the long-memory parameter we compute an implied value by calibrating the model with real data. We compare the performance of the three models using simulated data and we price options on the S&P 500 index.

1 Introduction

This article studies an integrated technique for option pricing, long-memory calibration, and parameter estimation, in stock and option markets with high frequency and long-range-dependent stochastic volatility. Before introducing some of the details of our study, we begin with a brief literature overview of stochastic volatility models and associated long-memory questions.

1.1 Stochastic volatility models

Stochastic volatility models were first introduced by Taylor [40, 41] and Hull and White [33] in order to account for inconsistencies in implied volatility values. More specifically, let us first imagine that a specific stock or index price $\{S_t\}$ truly follows the celebrated Black-Scholes model ([5], [37]), i.e. its stochastic dynamics are given by

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

where $\{W_t\}$ is a Wiener process (Brownian motion), and the volatility σ is a constant; Then the graph the volatilities which are responsible for the various call option prices observed on the option market as a function of the various strike prices, also known as the Black-Scholes "implied volatilities", would have to be a horizontal line at the level σ. However, it is well-known that such implied volatility graphs are hardly ever flat; it is much more common for them to look like a smile, or a smirk, meaning that certain far-from-the-money options have significantly higher implied volatilities than at- or near-the-money options. Some markets exhibit lower implied volatilities away from the money than near the money, in which one talks about an implied volatility frown.

In an effort to explain this phenomenon via mathematical modeling, both in discrete and continuous time, many authors have proposed that the volatility of the asset or index $\{S_t\}$ should be modeled as a random process itself. In other words, the observed process $\{S_t\}$ should follow the dynamics

$$dS_t = \mu S_t dt + \sigma(Y_t) S_t dW_t, \tag{1}$$

where $\{Y_t\}$ is the unobserved volatility process. Among the most popular continuous models are the Ornstein-Uhlenbeck mean-reverting process by Taylor [41] and Hull and White [33], and the CIR model introduced by Cox et al. [16]; widely used discrete-time models include the ARCH and GARCH time series models by Bollerslev [6], Bollerslev et al. [7] and Duan [21].

All these models have in common that they contain more sources of randomness than assets/indexes. Even in the case of single-asset models, the introduction of one extra source of randomness in the volatility makes the market incomplete, i.e. it is not necessarily possible to replicate any option within the market itself. In mathematical terms, this typically translates as the existence of more than one equivalent martingale measure for option pricing, and correspondingly more than one set of arbitrage-free prices for options. This problem is typically dealt with by referring to highly liquid options in order to help replicate other options. In principle, this requires being able to trade in the liquid option at high frequency. While this is an important practical problem, we do not deal with it in the present article: it turns out that our option prices are highly insensitive to the choice of a martingale measure; we refer to Chronopoulou and Viens [13] and Florescu and Viens [23] for a description of this phenomenon with the type of multinomial option pricing we use herein.

Another practical problem is that the asset's volatility process is never observed directly. This creates difficulties when it comes to parameter estimation. However, option pricing and statistical inference under stochastic volatility models have been extensively studied since their introduction. One can find an overview for option pricing techniques as well as parameter estimation procedures in the book by Fouque et al. [24] as well as a comparative review of various models in Taylor [42]. Again, high-frequency data is an important requirement for these procedures.

1.2 Long memory and stochastic volatility

As specialists in stochastic finance were coming to terms with the fact that stochastic volatility models may be required for sound modeling in many markets, a further modeling difficulty arose. Empirical studies were showing that some financial data exhibit long-range dependence as opposed to intermediate or short-range dependence. For general time series, long-range dependence, also called long memory, means that observations far apart in time are strongly correlated, as evidenced for instance by a very slowly decaying autocorrelation function. This is a slightly subtle phenomenon in financial data, since stock and index returns themselves are typically uncorrelated, while non-linear functions of the returns are correlated. Ding et al. [20] were among the first to observe that there is substantial correlation between absolute returns of the daily S&P 500 index prices. Evidence of long-range dependence came from the fact that fractional power transformations of the absolute returns exhibit high autocorrelations for high lags ([17], [19], [20]). Long term correlation was found using squared returns on various US indexes, in studies by de Lima and Crato [17], Lobato and Savin [36], and Breidt et al. [9]. Bollerslev and Mikkelsen [8] showed that the fractionally differenced absolute returns of the S&P 500 exhibit long-range dependence. Moreover, a slowly decaying autocorrelation function has also been observed in foreign exchange rates by Andersen and Bollerslev [2], and Henry and Payne [32].

Some of the first attempts to explain these slowly decaying autocorrelations were via long memory stochastic volatility modeling, and there now exists a wide variety of models for this purpose both in discrete and continuous time.

The first long memory stochastic volatility model was simultaneously introduced by Harvey [31] and Breidt et al. [9]. They suggested a discrete-time model in which the returns of the stock $\{X_t\}$ are described by

$$X_t = \sigma(Y_t)\epsilon_t,$$

where $\{\epsilon_t\}$ are iid shocks with zero mean and the logarithm of the volatility process $\{Y_t\}$ is described by a finite parameter *Fractional* ARIMA(p,d,q) model, that is

$$\phi(B)(1-B)^d Y_t = \theta(B)e_t, \tag{2}$$

where $\{e_t\}$ is a zero-mean serially uncorrelated process independent of $\{\epsilon_t\}$, $\phi(\cdot)$ and $\theta(\cdot)$ are polynomials in the lag operator B of orders p and q respectively and $d \in (-1/2, 1/2)$. This extension of the classical ARIMA(p,d,q) model for stochastic volatility, where d is no longer an integer, described the long-range behavior of the log-squared returns of market indexes successfully. Baillie et al. [3] worked in the same spirit, suggesting analogously to the fractional ARIMA(p,d,q) process of the mean, the *Fractionally Integrated* GARCH(p,d,q) process for $\{e_t\}$:

$$\alpha(B)(1-B)^d e_t^2 = \mu + (1-\beta(B))v_t, \;\; 0 < d < 1,$$

where $\{v_t\}$ are the innovations of the conditional variance and the polynomials $\alpha(B)$ and $(1-\beta(B))$ of orders p and q respectively, have all their roots lying outside the unit circle.

In continuous time, Comte and Renault [14] modeled the price process, $\{S_t\}$ as in (1) in which the dynamics of the volatility are described by a fractional Ornstein-Uhlenbeck process as follows

$$dY_t = \alpha(m - Y_t)dt + \beta dB_t^H,$$

where B_t^H is a fractional Brownian motion (fBm) with long-memory parameter $H \in (1/2, 1]$; fBm is the most basic continuous-time long memory process. Recently Comte et al. [15] extended the Heston option pricing model to a continuous time stochastic volatility model in which the volatility process is described by a square root long-memory process. In this way, they managed to describe volatilities with high persistence in the long run, without overincreasing the short run persistence. In contrast to the fractional Ornstein-Uhlenbeck model, the square-root one does not allow the volatility to attain negative values.

1.3 Pricing and statistical inference

Stochastic volatility models with long-memory either in discrete or continuous time are designed to describe important long-memory features of certain financial time series, which are not covered by standard stochastic volatility models. Their drawbacks include new difficulties in option pricing as well as in statistical inference. In this article we argue that these two issues are linked, and that in order to provide adequate solutions to these problems, it is essential to take advantage of all available information by using high-frequency data in an optimal way. We begin by summarizing some techniques recently developed in the literature, which we will use as tools and/or benchmarks.

Whether the underlying model is in discrete or continuous time, and indeed whether one believes that the underlying phenomena are discrete or continuous, one must still face the fact stock and index price processes are only observed in discrete time, and that volatility itself is never directly observed. The statistical inference problem of estimating volatility under these conditions is thus crucial from a financial modeling point of view, and is a stimulating question from the statistical standpoint. High-frequency quotes are then synonymous with high-quality information.

The key parameter to estimate turns out to be the long-memory parameter, also known as the Hurst parameter in honor of the hydrologist who first used fBm in scientific modeling, [35]. Comte and Renault, [14], suggested that one should begin by discretizing the model in order to obtain an approximate solution, and then use the *log-periodogram* regression approach to estimate the long-memory parameter. This regression technique was initially proposed by Geweke and Porter-Hudak, [28], known as the GPH estimator. A semiparametric modification of the GPH estimator was devised by Deo and Hurvich, [19], in the context of long-memory stochastic volatility models. In addition, Gao et al. [26, 27] proposed a methodology for estimating all parameters simultaneously for a discretized model. Casas and Gao, [10], studied the behavior of the (modified) GPH estimator for discretized fractional stochastic volatility models with an application in US financial indexes. More recently, Chronopoulou and Viens [13] showed how to compute an implied valued of the long-memory parameter H by calibrating to realized option prices (true prices observed on the option market).

Compared to the statistical inference literature, the problem of option pricing under long memory models has not been studied to the same extent. In continuous time, Comte and Renault, [14], adapted some of the most popular theories of bond pricing to long memory processes, under certain assumptions. Moreover, Chronopoulou and Viens, [13], proposed a multinomial recombining tree algorithm in which the volatility is sampled from its empirical distribution, which is estimated using a particle filtering algorithm. In discrete time, there are approaches by Bollerslev and Mikkelsen [8], Engle and Mustafa [22] and others mainly based on a time-varying conditional variance which is used either in the Black-Scholes context or in a simulation-based option pricing scheme (for instance the one by Amin [1]). The lack of interest in the option pricing question is inconsistent with the efforts devoted to understanding long memory in financial time series; it motivates us to pursue this research direction, and the closely related question of parameter estimation.

1.4 Outline of our results

In this article we consider three different long-memory stochastic volatility models. We begin with the one proposed by Comte and Renault, [14], in which the volatility is modeled by the continuous-time fractional Ornstein-Uhlenbeck process. We use the same option pricing and implied-H techniques developed in Chronopoulou and Viens [13]. More specifically, the algorithm of the methodology is summarized as follows:

1. For each value of H starting from 0.5 up to 0.95 we:

   (a) estimate the parameters of the model based on historical data,

   (b) run the particle filtering algorithm to compute the empirical distribution of the unobserved volatility,

   (c) use the multinomial recombining tree algorithm to compute the corresponding option prices for the specific value of H.

2. For each H, we compute the mean square error (MSE) of the calculated option price for specific strike prices from the center of the bid-ask spread. The bid or the ask price can also be used depending on whether we are interested in buying or selling an option.

3. The calibrated or *implied* value of H is the one that corresponds to the smaller MSE.

In addition, we propose an improved calibration procedure that takes into account the liquidity of the options, that is when we compute the MSE we use weighted option prices, in which the weights are proportional to the volume of the traded options.

Beyond option pricing, we wish to understand the effect of high-frequency data on the value of the implied H as well as on the other parameters of the model. Therefore, we compare high-frequency and non-high-frequency implied values of H for pricing options, with the interesting find that the empirical distribution of the volatility, and as a result the implied value of H, are not very sensitive to the data frequency. However, we obtain much more accurate estimates for the remaining parameters of the model with high-frequency data and as a result more accurate option prices.

In addition, it is natural to assume, and immediate to observe, that high-frequency is needed to track volatility properly. Although the volatility distribution is rather insensitive to the data frequency, unusual movements in the volatility (e.g during market shocks and announcement of economic indicators) can only be captured and tracked properly using high frequency data. Indeed, while tracking volatility in low frequency, pricing an option just after a large stock movement can result in a noticeable pricing error, since at that time the stochastic volatility filter is likely to be off the mark. It is not a surprise that H is largely the same regardless of the data frequency, since H measures volatility correlations over long lags, and may not be closely related to the rapidity of movements of the stock itself. The story would be different for a geometric fractional Brownian model (GfBm), since there the stock's movement amplitudes are more directly linked to H since (modulo the drift) the log returns are H-self-similar; but most researchers agree that, due to the presence of arbitrage and correlation of returns, GfBm is not appropriate for stock modeling.

The main disadvantage of the continuous time model is that the construction of the particle filter for the estimation of the empirical volatility distribution is computationally expensive. Therefore, we also study two discrete-time models. The first is the discretized version of the fractional OU model, using an Euler scheme, and the other is the fractional ARIMA model (2) proposed by Harvey [31] and Breidt et al. [9]. Using simulated data, we show that the required computational time to construct the volatility particle filter, compared to the continuous-time model, is significantly reduced in both discrete cases. We also find that our option-pricing methodology, adapted to both discrete models, works quite satisfactorily under simulated conditions. On the other hand, when using both discrete models to price options written on the S&P 500 index, we observe that the continuous-time model performs visibly better that the other two in this real-data situation; our criterion for performance is the ability to explain observed prices on the option market. This means that there is a non-trivial trade-off to be managed between computational expense and precision.

Furthermore, in order to compare the two discrete-time models, we study a model mis-specification problem. Assuming that the true model is the continuous time one, we numerically compare the performance of the two discrete-time models under our option pricing methodology. One would expect the discretized OU model to perform better than the Fractional ARIMA model, since the latter is not a discretization of the OU model, but one may wonder whether the option pricing problem is insensitive with respect to the type of discrete-time model one chooses. It turns out not to be: we show that the choice of discrete model makes a difference; the option prices computed using the discretized fractional OU model are a better match to those computed using the original continuous-time model. This observation is significant in practice, since we also determined that in the absence of computational constraints, option pricing under the continuous-time model is a better match to observed prices on the option market, and therefore discrete-time pricing models should seek to emulate it. The better performance of the discrete OU model is confirmed under our criterion of proximity to market prices.

The structure of this article is as follows. In Section 2, we study the continuous-time long memory stochastic volatility model and we discuss option pricing and parameter estimation techniques. We compare two different estimators for the parameter H and we investigate the effect of high-frequency data on the value of implied H, and on other parameter estimators. In Section 3, we introduce the discretized Ornstein-Uhlenbeck model and the discrete-time fractional ARIMA model; we adapt option pricing and estimation techniques and we compare the results for pricing options with these two models. In Section 4, we compare all three models by looking into their computational efficiency, model mis-specification issues, and their option-pricing performance with real S&P 500 index and option data. In the final section, we conclude our article with a summary and some recommendations.

2 Long memory stochastic volatility model in *continuous time*

2.1 The model; definitions.

The continuous time long memory stochastic volatility model (LMSV) we consider was initially introduced by Comte and Renault [14] and revisited by Chronopoulou and Viens [13]. If $\{X_t\}$ is the logarithm of the price process (dX_t is an infinitesimal log-return) and $\{Y_t\}$ the volatility process, then

$$\begin{cases} dX_t &= \left(\mu - \frac{\sigma^2(Y_t)}{2}\right) dt + \sigma(Y_t)\, dW_t, \\ dY_t &= \alpha\,(m - Y_t)\, dt + \beta\, dB_t^H, \end{cases} \tag{3}$$

where μ is the mean rate of return, α is the rate of mean reversion, m is the log-run mean of the volatility, β is the volatility of the volatility, σ is a chosen deterministic function, and $\{B_t^H\}$ is a fractional Brownian motion with Hurst index $H \in (0, 1]$.

Definition 1 *The fractional Brownian motion (fBm), $\{B_t^H\}$, with Hurst parameter $H \in (0, 1]$ is a centered Gaussian process whose paths are continuous with probability 1 and whose distribution is defined by its covariance structure:*

$$Cov(B_t^H, B_s^H) = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t - s|^{2H}).$$

Equivalently, the distribution is characterized by $B_0^H = 0$ *and* $Var\left[B_t^H - B_s^H\right] = |t-s|^{2H}$.

The parameter H characterizes both pathwise as well as distributional properties of the process and provides us with a classification according to its value: for $H < 1/2$ the process has rough paths and its increments exhibit medium-range dependence, while for $H > 1/2$ the paths, while still of infinite variation, are smoother, and its increments have long-range dependence. When $H = 1/2$, the process is the well-known standard Brownian motion (Wiener process), which has independent increments. The fBm is also H-selfsimilar and δ-Hölder continuous for any $\delta < H$. More details on fBm can be found in Nualart [39].

The main reason we choose to work with this process is its long-memory/ long-range dependence property, which we define as follows:

Definition 2 *A process* $\{X_m, m \in \mathbb{N}\}$ *is said to have long-range dependence (or long memory) if* $\sum_{n=1}^{\infty} \rho(n) = +\infty$, *where* $\rho(n)$ *is the autocorrelation function defined by* $\rho(n) = Cov(X_m, X_{m+n})/Var(X_m)$.

While the autocorrelation function ρ may depend on m, it does not when X is stationary, which is the case for the increments of fBm. Many non-stationary processes have auto-correlation functions whose dependence on m is weak enough that it does not effect the notion of memory length. We will not delve into these technicalities.

When $\sum_{n=1}^{\infty} \rho(n) < +\infty$, one often speaks of short-range dependence, although there are many scales of dependence within this class. Time series such as GARCH have exponentially decaying auto-correlation, which is truly short range, while the autocorrelation of fBm's increments with $H < 1/2$ decays like the power n^{2H-2}, which is much longer range than exponential decay, but still falls in the category of summable autocorrelation.

The volatility process $\{Y_t\}$ is the fractional analogue of an Ornstein-Uhlenbeck process. Thus, it is the unique process that satisfies the linear stochastic integral equation: $Y_t = \int_0^t \alpha(m - Y_s)ds + \beta B_t^H$, where α and β are drift and deviation parameters. The autocorrelation function of the increments of $\{Y_t\}$ inherits the long-range dependence property by fBm, when $H \in (1/2, 1)$, and is ergodic. The properties of this process have been extensively studied by Cheridito et al. [11].

The Ornstein-Uhlenbeck process is a popular model for standard stochastic volatility for many reasons, including the fact that it is mean-reverting. The same property holds true for the fractional Ornstein-Uhlenbeck process; the rate of mean reversion is α. An illustration of this fact is shown in Figure 1.

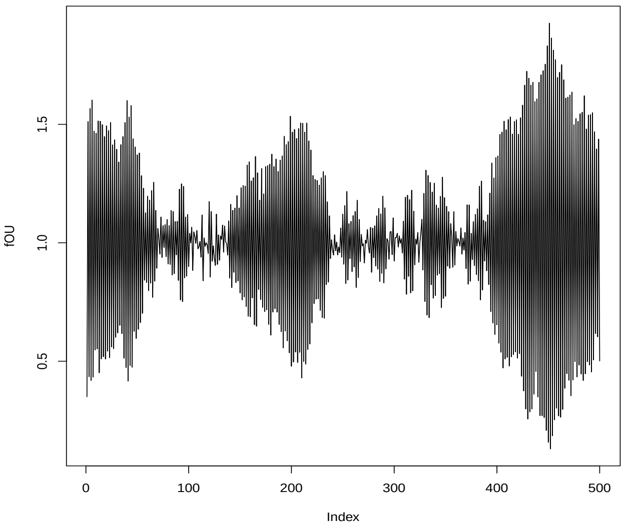


Figure 1: Fractional OU process with Hurst index 0.6 and mean reversion parameter $\alpha = 2$.

2.2 Option pricing

In this article, we choose to implement the option pricing scheme suggested by Chronopoulou and Viens [13]. We begin by describing the basic idea, referring to that article and references therein for the technical details, including an exhaustive description of the algorithm.

As mentioned in the introduction, our premise is that, although the basic LMSV model is in continuous time, we only have access to discrete time observations, namely the historical stock prices; the volatility process itself is not directly observed, even in discrete time. Our estimation and option-pricing methodology consists of two steps.

Step 1: *Estimation of the empirical distribution of the unobserved volatility.* This is handled by adjusting a genetic-type particle filtering algorithm by Del Moral et al., [18], and Florescu and Viens, [23]. This algorithm, using historical stock price observations, generates simulated pairs of stock and volatility values (the particles) one time-step into the future, and adjusts the probability weights of the particles based on their empirical likelihood when the next stock observation comes in. This can be considered as a non-parametric Bayesian approach, and is sometimes labeled as a *sequential Monte-Carlo algorithm* for computing the conditional distribution of the volatility given all past observations. The output is an empirical distribution for the unobserved volatility, the empirical measure of the weighted volatility particles. We call this the *(stochastic) volatility particle filter.*

Step 2 *Risk-neutral option pricing on a multinomial recombining tree*, constructed based on the full empirical volatility distribution produced by the previous step. This original technique was proposed by Florescu and Viens [23]; in each step of the tree, the value of the volatility is sampled from the volatility particle filter. The branches of the tree recombine unevenly in each step depending on the sampled value of the volatility, but the level of recombination is very high, typically close to binomial recombination. The tree construction is particularly faithful to the market's current volatility structure when used in high frequency. It also has the advantage of being computationally efficient, and closer in spirit to the way market makers compute option prices as a group, by focusing on current volatility beliefs based on past experience, rather than trying to incorporate theoretical volatility forecasts. The forecasting methodology is not uncommon in the SV option pricing literature (see Fouque et al. [24], e.g.), but does not produce more accurate prices (see Florescu and Viens [23]), and its implementation is typically much less efficient.

Since we work under a stochastic volatility model, it is important to determine the probability measure that we use for option pricing. Following the discussions in [13] and [23] for the quadrinomial (recombining) option pricing tree, if p is the probability of the upper (or lower branch), then the probabilities that correspond to the remaining branches are functions of p. However, it can be shown that p is restricted to the interval $[\frac{1}{12}, \frac{1}{6}]$. If we plot the computed prices for values of p varying from $\frac{1}{12}$ to $\frac{1}{6}$, we observe that the option price is quite insensitive to the choice of p; this is also illustrated in Figure 2.

Remark 1 *The methodology we propose for option pricing can also be used when the volatility process is described by a fractional square-root process as in [15] (or any other mean reverting long memory process). The same holds for the implied H procedure. However, the parameter estimation technique as described in the following section is restricted only to the fractional OU model, since it is based on its specific characteristics and properties.*

2.3 Long memory calibration; S&P 500 data.

When the underlying model is the continuous-time LMSV model we have to estimate the following parameters: the drift μ, the rate of mean reversion α, the long-run mean m, the volatility of the volatility β, and the long-memory parameter H. As mentioned by the authors in [13], proper estimation the long-memory parameter H is of the utmost importance, since

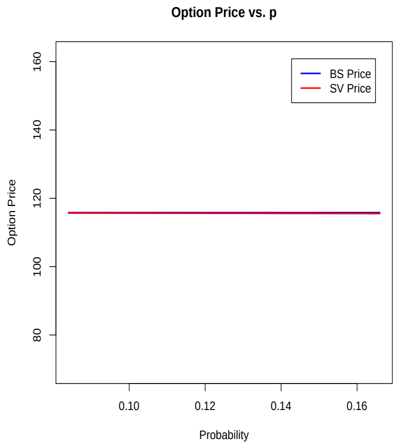


Figure 2: Option prices for different choices of the free probability $p \in (\frac{1}{12}, \frac{1}{6})$ and comparison with the corresponding prices computed based on a constant volatility Black-scholes model.

it significantly affects the estimated empirical volatility distribution as well as the remaining parameters of the fractional diffusion, and thus weighs heavily on computed option prices.

We propose to determine H by calibrating the model using realized option prices. We therefore obtain an implied value for H. The calibration method is simple, and presents itself naturally in our option-pricing context. The main idea is to repeat the following procedure for values of H varying from 0.5 to 0.95 with a rather fine step (e.g. 0.01):

Integrated estimation, pricing, and calibration procedure

- For each fixed value of H, we start by estimating the parameters of the model based on historical data. We can do so by modifying standard techniques based on the variogram, as described in the book by Fouque et al. (Chapters 3,4, [24]). In this step, as discussed in detail in the following section, it is crucial to use high-frequency data; in this way, the produced estimates will be consistent.
- Once we obtain the estimates for all the parameters of the model, we construct the volatility particle filter, which is the first step of the pricing algorithm.
- Then, we move to the second step of the pricing algorithm and we price options for different strike prices, K, for a certain maturity, T.
- The last step consists of computing the mean-square error (MSE) of the computed option prices with the center of the bid-ask spread for the corresponding options from the market.

Criterion	**Implied $\hat{\mathbf{H}}$**
Bid	0.55
Ask	0.52
Center = (Bid+Ask)/2	0.53

Table 1: Choice of Implied H when we calibrate the model with the *bid*, the *ask* or the *center* of the bid-ask spread. (The results are based options written on the S&P 500 index on March 30^{th} 2009 with maturity $T = 35$ days).

- The implied value of H is the one that corresponds to the smaller MSE.

In order to empirically test the stability of our "estimator" for H, we repeat this procedure 1000 times for the same data and then we average all the implied values of H. This implied H is the value of the Hurst parameter that we are going to work with in the future for option pricing. Moreover, this Monte-Carlo type estimate gives us an empirical variance for the proposed estimator. Using this value of H as fixed, we can now compute the final estimates for remaining parameters of the model. Since the use of high frequency data for their proper estimation is crucial, we discuss the details of this procedure in the following section.

One question that arises from a practitioner's point of view is why choose the center of the bid-ask spread and not the bid or the ask price in our criterion for choosing H. The center of the bid-ask spread has the advantage that we can price in a universal way a wider range of options (e.g. for all strike prices). However, it is interesting to investigate the effect on implied H if the comparison is done with respect to the bid or the ask price.

We consider a real data example: we price a European call written on the S&P 500 index on March 30th, 2009, expiring in 35 business days. The interest rate during this period is $r = 0.21\%$ and the index at the time of pricing is worth $S_0 = \$787.53$. As shown in Table 1 there are small differences in the choices of H. Therefore, depending on the type of options that we are interested in pricing we can choose one criterion over another.

Similarly, we can concentrate on a specific range of options in computing the implied value of H. More specifically, if we are interested in in-the-money call options we can repeat the analysis only for strike prices below S_0. For the same real-data example, as it is shown in Table 2, the computed values of H are different, however they match the realized prices much better and the improvement can be seen in Table 3.

Before comparing the implied H with another popular estimator of H in the literature, we propose a modification of the calibration procedure, based on the volume of trades for each strike price. In order to ensure that the estimate for H reflects the general consensus of the

Range of Ks	**Implied $\hat{\mathbf{H}}$**
\$670 - \$740	0.54
\$750 - \$800	0.53
\$810 - \$850	0.54

Table 2: Values of implied H for particular ranges of K. The stock price 'today' is S_0=\$787.53. (The results are based options written on the S&P 500 index on March 30^{th} 2009 with maturity $T = 35$ days).

Strike Price	Bid	Ask	**Implied $\hat{\mathbf{H}}$** $= 0.53$	**Implied $\hat{\mathbf{H}}$** $= 0.53$ or 0.54
670	126.9	130.3	123.96	127.92
680	118.5	121.9	115.81	119.65
690	110.4	113.8	107.93	111.22
700	102.6	105.9	100.32	104.35
710	94.6	98	92.00	92.68
720	87.1	90.5	85.97	89.66
730	79.8	83.2	79.26	80.02
740	73	76.2	72.86	74.50
750	66	69.5	66.80	66.80
760	59.7	63	61.07	61.07
770	53.5	57	55.66	55.66
780	47.8	51.3	50.59	50.59
790	42.6	45.7	45.85	45.85
800	37.4	40.8	41.43	41.43
810	32.8	36.2	37.33	36.05
820	28.3	31	33.54	29.35
830	24.6	27.9	30.04	26.33
840	20.8	24.3	26.83	21.05
850	18.9	20.8	23.90	19.66

Table 3: Computed European call option prices on the S&P 500 using a universal value of implied H or the local values of implied H as shown in Table 2

Strike Price	Bid	Ask	**Weighted Implied** $\hat{\mathbf{H}} = 0.534$
670	126.9	130.3	125.23
680	118.5	121.9	118.0
690	110.4	113.8	110.33
700	102.6	105.9	103.25
710	94.6	98	94.75
720	87.1	90.5	87.04
730	79.8	83.2	78.32
740	73	76.2	74.21
750	66	69.5	66.54
760	59.7	63	61.25
770	53.5	57	54.36
780	47.8	51.3	49.96
790	42.6	45.7	45.24
800	37.4	40.8	40.32
810	32.8	36.2	35.23
820	28.3	31	31.45
830	24.6	27.9	28.45
840	20.8	24.3	25.33
850	18.9	20.8	21.65

Table 4: Computed European call option prices on the S&P 500 using a weighted implied H.

market we compute the MSE by using weighted option prices with weights that are proportional to the volume of the trades. More specifically, we take the weight that corresponds to the i^{th} strike price to be

$$w_i = \frac{\#\text{ of trades at strike } K_i}{\text{Total } \#\text{ of trades}}.$$

The weighted option prices for the previous example are summarized in Table 4 and a comparison with the un-weighted prices is illustrated in Figure 3. The weighted option prices stay truer to the bid-ask spread, even with far-from-the-money calls.

Remark 2 *One issue that arises here is the time during which the value of H is considered to be constant. Empirical evidence shows that we are not able to consider H to be constant for a period of 2 or 3 years or more. H seems to be constant for a period of less than a year, depending on the stability of the market.*

Figure 3: Comparison of computed option prices based on the weighted and the un-weighted calibration procedure.

When major market shocks occur, such as the financial "meltdown" of October 2008, H is no different than many other economic and financial indicators, in the sense that it is expected to change abruptly. H would typically increase in the weeks after such an event, as the memory of the crash remains strong in the market makers' actions: see Chronopoulou and Viens [13].

The rationale behind the calibrating of H to the option market is that among all financial professionals, none think harder and more frequently about short and medium term volatility forecasting (days to months) than the market makers in liquid option markets. Our implied H method allows us to distill the long memory parameter out of these market makers' activity.

Long-memory parameter estimation is notoriously treacherous with financial data. One reason is because many popular methods use self-similarity estimators, even though the connection to long memory, which is well-known for models such as fBm, is not appropriate for LMSV because the volatility is not directly observed, and the OU process is not self-similar, strictly speaking. Another reason may be that long memory discrete time series are often difficult to fit to financial data ([12]). Among the most popular Hurst index estimators for the LMSV models in the literature is the log-periodogram regression estimator, also known as the GPH estimator, initially introduced by Geweke and Porter-Hudak [28]. This estimator is mainly based on the discretization of the model and then on the application of a maximum likelihood method on the spectral domain, by minimizing the Whittle contrast function (an approximation of the log-likelihood function). More details regarding the GPH estimator can be found in Casas and Gao [10] and Geweke and Porter-Hudak [28]; a discussion regarding this estimator and the implied H can be found in our article [13].

To illustrate this discussion, we numerically compare the GPH estimator and our proposed method of finding an implied value of H by calibrating it to option prices. We use S&P 500 data during three different periods: April 2008, May 2008 and March 2009. In all cases we consider two months of historical data in order to estimate H. To compute the implied H we generate filters of $n = 1000$ particles, using $M = 10000$ Euler steps for the simulation of the model, and $N = 100$ tree-steps in the multinomial tree algorithm. Using the generated volatility particle filters, we price call options which we compare with the center of the corresponding bid-ask spread for market prices. The GPH estimator is computed using the same historical data. The results we obtain are summarized in Table 5.

From Table 5, we observe that the two methods produce significantly different values of H. In the case of May 2008, the GPH estimator is far below $1/2$; such a medium memory situation could be an indication of anti-persistence, which can be interpreted as an extremely high rate of mean reversion rather than of memory length. We are not aware of any works

"Today"	**Implied $\hat{\mathbf{H}}$**	Empirical Std. Dev.	**GPH $\hat{\mathbf{H}}$**	Empirical Std. Dev.
April 4th 2008	0.51	0.0035	0.66	0.022
May 2nd 2008	0.50	0.0011	0.23	0.045
March 30th 2009	0.53	0.0201	0.67	0.0098

Table 5: Comparison of the estimated long-memory parameters using the implied H technique and the Geweke and Porter-Hudak method.

in which either this or the memory length interpretations are believed to be accurate in the case of liquid and efficient option markets such as the one we study here. For the same month of May, our implied H method gives an H equal to 1/2, meaning that there is no noticeable memory, while there was some memory detected a month prior, in April. This is consistent with the remarks made in Chronopoulou and Viens [13] regarding the effect on the markets of the March 2008 collapse of Bear Stearns, the smallest of the "big five" Wall Street independent investment banks.

To further investigate and compare the two estimators, we feed both values of H into our option pricing algorithm for a European call written on the S&P 500 on March 30th, 2009, expiring in 35 business days. The interest rate during this period is $r = 0.21\%$ and the stock at the time of pricing is worth $S_0 = \$787.53$. We compare our option prices with the bid-ask spread on the option market on that day for strike prices varying from $K = \$670$ to $K = \$850$. These results are summarized in Table 6. Because of the nature of our implied method, the option prices computed using the implied H would have to be at least as close to the prices realized on the option market, as those computed with the GPH estimator. In Table 6 we can see the detail of how much closer option prices computed with an implied H are to the market call prices, than those computed with the GPH estimator. The difference is quite significant near the money and out of the money, while neither estimators perform well for in-the-money call options, with the GHP estimator actually doing a bit better than the implied H one. Recalibrating H for the in-the-money range alone, as described previously, would address this deficiency; such a procedure is not available for the GHP estimator, since it is not based on option prices.

2.4 Statistical inference and high-frequency S&P 500 data

So far, we have only discussed and analyzed methods for estimating the long-memory parameter. However, the model contains several other parameters that also need to be estimated as accurately as possible.

Strike Price	Bid	Ask	**Implied** $\hat{\mathbf{H}} = 0.53$	**GPH** $\hat{\mathbf{H}} = 0.67$
670	126.9	130.3	123.96	128.54
680	118.5	121.9	115.819	120.82
690	110.4	113.8	107.935	113.89
700	102.6	105.9	100.324	106.90
710	94.6	98	92.9996	99.93
720	87.1	90.5	85.9752	92.99
730	79.8	83.2	79.2616	86.02
740	73	76.2	72.8693	79.054
750	66	69.5	66.805	72.614
760	59.7	63	61.0704	67.95
770	53.5	57	55.6668	63.324
780	47.8	51.3	50.5955	58.602
790	42.6	45.7	45.8526	53.961
800	37.4	40.8	41.4336	49.295
810	32.8	36.2	37.3349	44.660
820	28.3	31	33.5426	40.023
830	24.6	27.9	30.0468	35.478
840	20.8	24.3	26.8366	33.070
850	18.9	20.8	23.9021	30.581

Table 6: Computed European call option prices on the S&P 500 using the algorithm described in Section 2.1, using two different values of H; the Implied and the GPH.

For each fixed value of H, we can estimate the other parameters by following standard approaches. More specifically, we consider a variogram analysis in order to obtain estimates for the rate of mean reversion α and for the volatility of the volatility β. More details regarding this approach can be found in the book and a related article by Fouque et al. [24, 25]. Here, we briefly discuss the modifications of this procedure in the case of a fractional Ornstein-Uhlenbeck process.

Assume that we have access to high-frequency observations (at least once every 5 minutes, say). Let X_n denote the nth five-minute average of the price at time $t_n = n\Delta t$, where $\Delta t = 5$. Then consider the fluctuation of the data

$$D_n = \frac{2(X_n - X_{n-1})}{\sqrt{\Delta t}(X_n + X_{n-1})},$$

or in other words the observed realization of the asset price return. Model D_n as $D_n = \sigma(Y_n)\epsilon_n$, where ϵ_n is a sequence of iid zero-mean, variance-one, random variables, with ϵ_n independent of Y_n, and let L_n be its logarithm, $L_n = \log|D_n|$. The ϵ_n's are also chosen so that $\mathbf{E}(\log|\epsilon_n|) = 0$. This model is consistent with the volatility part of our LMSV semimartingale model (3), with σ a predetermined non-random function. The drift part of our LMSV will be negligible in the variogram's high-frequency asymptotics, just as the drift part of a semimartingale does not appear in its quadratic variation expression. Therefore, the model for D_n is consistent with our LMSV model, as long as Δt is small enough, i.e. as long as the frequency is high enough.

To be specific, the variogram of L_n is defined as

$$V_j^N = \frac{1}{N}\sum_{n=1}^{N}(L_{n+j} - L_n)^2,$$

where j is the lag and N the total number of points. Then, we have that

$$\mathbf{E}(L_{n+j} - L_n)^2 = \mathbf{E}(L_{n+j}^2) + \mathbf{E}(L_n^2) - 2\mathbf{E}(L_{n+j}L_n)$$

We compute each term separately as follows:

$$\begin{aligned}
\mathbf{E}(L_n^2) &= \mathbf{E}([\log\sigma(Y_n) + \log|\epsilon_n|]^2) \\
&= \mathbf{E}(\log\sigma(Y_n)^2) + \mathbf{E}(\log|\epsilon_n|^2) + 2\mathbf{E}(\log\sigma(Y_n)\ \log|\epsilon_n|) \\
&= \nu_{\sigma,H}^2 + c^2,
\end{aligned}$$

where $c^2 = \mathbf{E}(\log|\epsilon_n|^2)$ and $\mathbf{E}(\log\sigma(Y_n)\ \log|\epsilon_n|) = 0$, and asymptotically for small Δt,

$$\nu_{\sigma,H}^2 := \nu_{\sigma,H}^2\,(\alpha,\beta) = H\,\Gamma(2H)\,\beta^2 + H\,\Gamma(2H)\,\beta^2\alpha^{1-2H} - \frac{\beta^2}{2}j^{2H} + o(j^{2H}).$$

Moreover, we have

$$\mathbf{E}(L_n\, L_{n+j}) = \mathbf{E}(\log\sigma(Y_n)\log\sigma(Y_{n+j})) = \nu^2_{\sigma,H}\; e^{-\alpha\, j\, \Delta t}$$

Therefore, we finally obtain

$$\mathbf{E}(L_{n+j} - L_n)^2 \approx 2c^2 + 2\nu^2_{\sigma,H}\,(\alpha,\beta)\,(1 - e^{-\alpha\, j\, \Delta t}).$$

Thus, we can use standard regression to estimate the parameters of the following linear equation

$$V_j^N = 2c^2 + 2\nu^2_{\sigma,H}\,(\alpha,\beta)\,(1 - e^{-\alpha\, j\, \Delta t}),\;\; j \in \mathbb{N}.$$

and consequently we can extract the desired estimates for α and β given that H is known.

In the case of the exponential fractional OU model, in which we take $\sigma(x) = e^x$, we can estimate the quantity $s^2 := \mathbf{E}(D_n^2)$ by the mean square of the fluctuations D_n. However, it is known that $s^2 = e^{2m+\nu^2}$, where we can take $\nu^2 = \frac{\beta}{2\alpha}$.

In the literature when it comes to statistical inference for stochastic volatility models, it is suggested, if not required, to use high frequency data, because by doing so we obtain consistent estimators, and, as we explained above, we also simplify the estimation procedure by decoupling drift and volatility parameters. Therefore, if the value of H is known, then the procedure above dictates for us to use high-frequency data for the estimation of the remaining parameters. However, there is nothing in our implied methodology for estimating H that says it should also be done using high frequency data. What is the effect of high-frequency data on the empirical distribution of the volatility (i.e. volatility particle filter)? And as a result, what is the effect on our choice of implied H? It is a simple matter to investigate this question in practice.

Using the March 2009 S&P 500 intraday (high-frequency) data for a period of one month, we compute the implied value of H. The value we obtain is the same (i.e. $H = 0.53$) as with the use of low frequency (daily) data. More generally, for the purpose of long-memory calibration, our volatility filter performs well even when the data have lower frequencies. Therefore, our recommendation is to use low frequency data for the particle filtering algorithm, but high-frequency data for the statistical estimation of the remaining parameters of the model. This greatly increases the efficiency of the filtering algorithm for the purpose of estimating H.

An explanation for the relative insensitivity of the implied estimation H to the data frequency, lies in the fact that in the case of long memory, the auto-correlations decrease slowly, and therefore an interpolation of their behavior using few data points yields essentially the same amount of information as a method which tracks the volatility much more closely. While this robustness property holds for calibrating H in the context of option pricing, those who wish to track volatility closely by using the particle filter (a task for which particle filtering, a.k.a. sequential Monte-Carlo methods, was designed), will still need to use high-frequency data.

2.5 Computational difficulties with the continuous time LMSV model

To summarize, using the implied-H approach, we are able to overcome the obstacles of an unobserved volatility process with long-memory and extract the appropriate value of H that provides us option prices which are close market prices.

However, there is a hidden disadvantage in our approach: the computational time that is needed to construct the volatility particle filters is high, even when using relatively low-frequency data. We need to take into account that in order to better simulate the continuous-time fractional diffusion, we need – on average – more than 500 Euler steps. As it turns out, the smaller the value of H the more Euler steps we need. This is because the speed of convergence of the Euler scheme is proportional to n^{-H} ([38]). To make matters worse, it was noted in Chronopoulou and Viens [13] that the higher the value of H, the more particles are need in order to have a satisfying range of volatility values, i.e. to avoid filtering with too few particles. This effect is presumably due to the fact that for high H, the volatility process is too "sticky", and does not explore a sufficiently large area of its state space in short periods of time to rely on the same number of particles as in cases that are closer to no memory (H closer to 1/2). As a result, by increasing both the numbers of Euler steps and of particles, we also increase the computational time needed to generate the particle filter. All these problems are compounded by our need to generate separate filters for a number of different values of H, in order to implement the implied-H method.

3 Long-memory stochastic volatility models in discrete time

To overcome the computational disadvantage described in the last paragraph above, which affects the continuous-time LMSV model in its use for computing the implied value of H, we now turn to the discrete time world. The first model we propose is the discretized version of the continuous-time model. The second one, while not linked to the LMSV model in a strong statistical sense, is nonetheless a discrete time analogue of it, one which is commonly used by financial econometricians, one in which the volatility is a fractional ARIMA process.

3.1 Discrete time models

3.1.1 Discretized fractional OU model

The first discrete time model we consider is the discretized fractional Ornstein-Uhlenbeck model. We use an Euler discretization scheme for both equations, thus we have

$$\begin{cases} X_n &= X_{n-1} + \left(r - \frac{\sigma(Y_{n-1})^2}{2}\right)\delta_n + \sigma(Y_{n-1})\sqrt{\delta_n}\epsilon_n \\ Y_n &= Y_{n-1} + \alpha(m - Y_{n-1})\delta_n + \beta(B_n^H - B_{n-1}^H), \end{cases}$$

where ϵ_n are zero-mean iid Gaussian random variables and B^H is a fractional Brownian motion. The discretized version of the fractional equation inherits the desired property of long-range dependence from the increments $B_n^H - B_{n-1}^H$ of fBm, which are commonly known as *fractional Gaussian noise* (as opposed to the ubiquitous Gaussian *white noise*). The rate of convergence of the Euler scheme for the Geometric Brownian Motion (the process X_n) is of order $n^{-1/2}$, while the rate of convergence of the fractional stochastic differential equation (the process Y_n) is of order n^{-H}. This implies that for smaller values of H, more observations are needed in order to approach the continuous time model. More details regarding the discretization of the fractional stochastic differential equation can be found in [38].

3.1.2 Fractional ARIMA model

The *fractional ARIMA* model was proposed by Harvey [31] and Breidt et al. [9] simultaneously, and was among the first LMSV models introduced in the literature. These authors propose describing the volatility by a fractional ARIMA$(1, d, 1)$ process. More generally, the fractional ARIMA(p, d, q) model for volatility is as follows.

Definition 3 *Let $\phi(\cdot)$ and $\theta(\cdot)$ be polynomials of orders p and q respectively. Let Y_n be a stationary process such that*

$$\phi(B)(1 - B)^d Y_n = \theta(B)e_n,$$

for some $d \in (-1/2, 1/2)$ and a sequence $(e_n)_n$ of iid variables with mean 0 and variance 1. Then, the process Y_t is called a fractional ARIMA(p, d, q) process.

Definition 4 *More specifically, to model financial volatility, with $H = d + 1/2$, we may use a fractional ARIMA(p, d, q) process Y as above, and define the returns of the stock or index as*

$$X_n = \sigma\left(Y_n/2\right)\epsilon_n$$

where σ is a predetermined function, and $(\epsilon_n)_n$ is another sequence of iid variables with mean 0 and variance 1.

The Fractional ARIMA(p, d, q) model is an extension of the classical ARIMA(p, d, q) model with integer d, first introduced by Granger and Joyeux [29].

The parameter $d = H - 1/2$ describes the memory of the underlying model. The fractional ARIMA model is long-range dependent when $d \in (0, \frac{1}{2})$. More details regarding these models can be found in Beran [4].

Remark 3 *The above defines the fractional ARIMA in the general case. Part of the problem for such a model is the procedure in order to specify the parameters p and q. For this, we refer to Hamilton [30] and to standard time series analysis techniques. In the sequel, we choose to work with a Fractional ARIMA*$(1, d, 1)$ *model: it was suggested to be the most appropriate for modeling S&P 500 data by [3]. We also choose to use Gaussian error sequences* $(\epsilon_n)_n$ *and* $(e_n)_n$.

3.2 Option pricing

Let us now explain how to adjust the option pricing algorithm developed in [13] and expanded in Section 2 herein, for the case of discrete time. The change of the underlying model affects only the algorithm for the construction of the empirical distribution of the volatility process via the particle filter. The construction of the tree, and the pricing algorithm, remain unchanged. In the particle filtering algorithm, instead of simulating the continuous time model using an Euler technique over one observation time step, we simply simulate the value of the discrete-time model in a single go, over this time step. This significantly reduces the computational time of the particle filtering algorithm. We formalize this simple modification.

Remark 4 *Referring to Section 3.1 of [13], the following changes are sufficient. Steps (1), (2) and (3) of the Mutation Step are replaced by the simulation of the pairs* $\{X^j_{t_i},\ Y^j_{t_i}\}_{j=1,\dots,n}$ *at time* t_i *from time* t_{i-1}*, for* $i = 1, \dots, K$*, using the definition of the discrete-time model. Here* n *is the number of particles used in the algorithm and* K *is the number of (daily) observations that we have available. Thus, to obtain each such pair at time* t_i*, we no longer need* M *Euler steps, but one single simulation.*

We test the performance of our approach under both discrete time models using simulated data: We simulated each model separately and we then compute call option prices for a call option that expires in 35 days with parameters: $S_0 = \$800$ and $r = 0.21\%$, and compare our values to those using a binomial model fitted to the historical volatility $\sigma_{hist} = 0.26$ given by both our discrete simulation. This comparison permits us to get an idea of how far are our

Strike Price	Discretized Model	fARIMA Model	Binomial Model
670	131.07	131.14	131.04
680	121.53	121.29	121.48
690	112.13	112.04	112.05
700	103.30	102.56	102.79
710	93.314	93.904	93.847
720	86.764	84.931	85.177
730	77.662	76.975	76.791
740	69.330	69.125	68.756
750	61.941	61.473	61.261
760	55.941	54.274	54.146
770	47.709	49.014	47.399
780	43.209	40.453	41.396
790	35.582	34.892	35.799
800	30.823	30.992	30.582
810	25.078	26.857	26.159
820	21.661	23.124	22.087
830	18.379	17.212	18.437
840	13.280	14.644	15.419
850	12.948	13.562	12.664

Table 7: Call Option Prices for a simulated model using both discretized and fARIMA models and comparison with the corresponding classical Binomial prices. The Call option that we priced has the following parameters: $S_0 = \$800$, $\sigma_{hist} = 0.26$, $r = 0.21\%$ and $T = 35$ days. (Note: both simulated models had $\sigma_{hist} = 0.26$.)

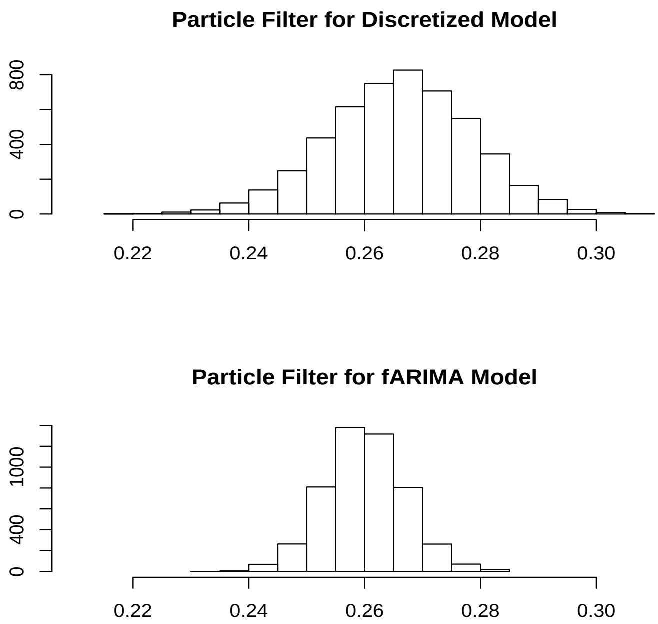


Figure 4: Particle Filters.

results are from a standard model. The results of the simulation experiment are summarized in Table 7 and Figure 4.

From Figure 4 we can say that the particle filtering algorithm works satisfactorily when the underlying model is a discretized fractional OU or a fractional ARIMA model. The option prices we compute, summarized in Table 7, are reasonable in all cases.

3.3 Statistical inference

Depending on which discrete-time model we use, estimation procedures specific to each model are used for all parameters, except H (or equivalently d), which is estimated using the same implied-H approach in both cases.

For the discretized version of the fractional Ornstein-Uhlenbeck model, we propose using the variogram approach described in Section 2.4.

For the fractional ARIMA model, we follow the standard parametric approach, as described by Breidt et al. [9]. Therefore, we choose to work with the logarithms of the squared returns, i.e. $Z_t = \log X_t^2$, which have a signal-plus-noise representation as follows

$$Z_t = \mu + Y_t + \eta_t,$$

where $\mu = \mathbf{E}[\log \epsilon_t^2]$ and $\eta_t = \{\log \epsilon_t^2 - \mathbf{E}[\log \epsilon_t^2]\}$. Traditional parameter estimation techniques developed for other autoregressive stochastic volatility models fail in this case, because of the long-memory property.

Breidt et al. [9] propose to estimate the parameters of the model from $\{Z_t\}$ using the Whittle approximation to the likelihood function. Given the observations $z_t = \log X_t^2$, for $t = 1, \dots, K$, define the periodogram

$$I_j = \frac{1}{2\pi K} \left| \sum_{t=1}^{K} z_t \exp(-i\, \omega_j\, t) \right|^2, \quad j = 1, \dots, K-1$$

where $\omega_j = 2\pi j/K$ are the Fourier frequencies. The Whittle approximation for the -2 log-likelihood is

$$L_W(\theta) = \sum_{j=1}^{\lfloor (K-1)/2 \rfloor} \left[\log f_{Z,\theta}(\omega_j) + \frac{I_j}{f_{Z,\theta}(\omega_j)} \right]$$

where $f_{Z,\theta}(\omega_j)$ is the spectral density for Z at frequency ω_j under the parameter model θ. The *spectral density function* is defined as the inverse Fourier transform of the autocovariance function as follows

$$f(\lambda) := \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-i\lambda h} \rho(h).$$

The authors in [9] also established the consistency of the Whittle estimator.

4 Comparison between *discrete* and *continuous* time models

Continuous and discrete time models have fundamental differences. The use of one class of models over the other can depend on the underlying phenomenon, but this determination cannot be made clearly when the underlying physical phenomenon may be in continuous time, while one only has access to discrete time observations, and these observations cannot be obtained with arbitrarily high frequency. In this case, both types of models can be used, and each one has its advantages and disadvantages.

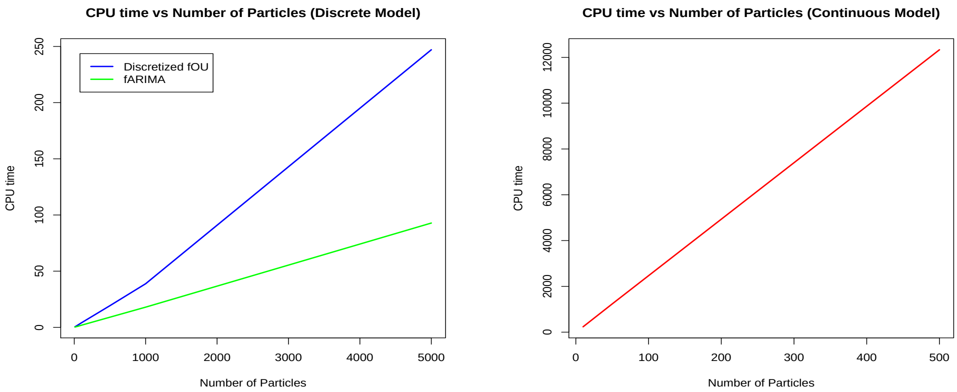
Discrete time models are easier to handle and typically require less complicated tools for statistical inference. However, they tend to be very sensitive to the frequency of the data. The models that corresponds to tick or daily data are far from identical, and are also different from the one that corresponds to weekly or monthly observations. Continuous time models do not depend on the frequency of observations, but their parameter estimation is more involved; in particular, when observations are in discrete time, either the model must be discretized, to allow for discrete time techniques for parameter estimation, or one must derive continuous-time estimators which one then must discretize in order to evaluate them.

In our study, the continuous time model has one additional disadvantage: it is computationally expensive when implemented with real data. However, that the underlying volatility phenomena are in continuous time, constitutes a legitimate belief particularly in the martingale modeling world of option pricing, so that the continuous-time model may be the most appropriate one. In this section we provide empirical evidence that this is correct, by numerically comparing the three models discussed in the previous sections and using our S&P 500 example. We compare the CPU time required for the generation of volatility particle filters, study a question of misspecification, and evaluate the models' performances in option pricing.

4.1 Reduction of CPU time

In this section, we compare the CPU time that each model needs in order to construct one filter, for various different numbers of particles. We prefer to use simulated data for the comparison, to eliminate the uncertainty on estimated parameters.

We simulate each model and use these simulated data as "historical prices" which we feed into the particle filtering algorithm. We then compute the corresponding CPU time for the number of particles varying from 10 to 5000. We use 255 data points (corresponding to one year daily data) and choose $H = 0.53$ or equivalently $d = 0.03$, for the long-memory parameter.



(i) (ii)

Figure 5: (i) CPU time needed for the continuous time model, (ii) CPU time needed for the discrete time models.

The results are summarized in Figure 5. We observe that the Fractional ARIMA model is the fastest one, although the difference with the discretized OU model is not significant, because the real time as perceived by the user is trivial in both cases. It is not possible to include the continuous-time filter CPU time on the same graph, since the ratio with discrete-time performance is 2 orders of magnitude higher. We conclude that our initial goal to reduce computational time by moving to discrete time models, is achieved.

4.2 Model mis-specification study

As discussed before, even though we have access only to discrete time observations, it is legitimate to believe that the true phenomena occur in continuous time. Therefore, in this section, we wish to discuss a model mis-specification issue. Assume that the continuous-time LMSV model is the true underlying model for stock movements. This means that using any discrete-time model constitutes a model mis-specification. How significant is this problem, in each of the two discrete models we study?

We study the effect of this mis-specification on the empirical distribution of the unobserved process first, and then on option prices. We simulate a continuous fractional OU model using the following parameters: $\alpha = -0.02733$, $\beta = 0.07567$, $\mu = -0.00138$ and long-memory parameter $H = 0.6$.

We assume we have historical observations for $K = 100$ trading days and we also simulate intraday data. Note that in all cases we assume that H is known. Using the techniques discussed in Section 2.3, the estimated parameters of the discretized fractional OU model are found to be:

$$\alpha = -0.02633, \quad \beta = 0.06673, \quad \mu = -0.00098.$$

Using the approach in Section 3.2, we fit a fractional ARIMA$(1; 0.1; 1)$ model with estimated parameters

$$\phi = 0.9831, \quad \theta = -0.0025.$$

For each of the three models we estimate the empirical distribution of the unobserved volatility using $n = 1000$ particles. For the continuous time model we use 10000 Euler steps. The three histograms can be seen in Figure 6. We observe that using discrete and continuous time OU models we construct very similar volatility filters, in terms of the range of the volatility values and the corresponding probability weights. On the other hand, the fractional ARIMA model gives higher weights to lower volatility values. Since our data is actually simulated, it is possible to see how well our filters track the true volatility. The ARIMA model misspecifies the volatility much more significantly than the discretized OU model, even in the relatively moderate daily data frequency.

Using the empirical distribution of the volatility, for each model, we price a European Call option with the following parameters: stock price today $S_0 = 120$, interest rate $r = 0.10\%$, time to maturity $T = 100$ trading days and strike prices varying from $K = 90$ to $K = 210$. The results are summarized in Table 8. Again, the two OU models gives us option prices that are close to each other, while the fractional ARIMA model is further afield, particularly for large strike prices.

We conclude that in a continuous-time world, model mis-specification due to discretization is more serious for the fractional ARIMA model than for the discretized OU model, even with moderate-frequency (daily) data.

4.3 Comparison on S&P 500 data

In this final section, we compare the three models using real data. We focus on the same example as in Section 2.3: to price a European call option on the S&P 500 index on March 30th, 2009 with time to maturity $T = 35$ business days. The first step is to compute the implied-H value for each model. In Section 2.3 this was found to be $H = 0.53$ for the fractional OU model. It turns out that we obtain the same implied value of H when we use the discretized fractional OU model, while using the fractional ARIMA $(1, d, 1)$ model we obtain $H = 0.52$. The results are repeated in Table 9.

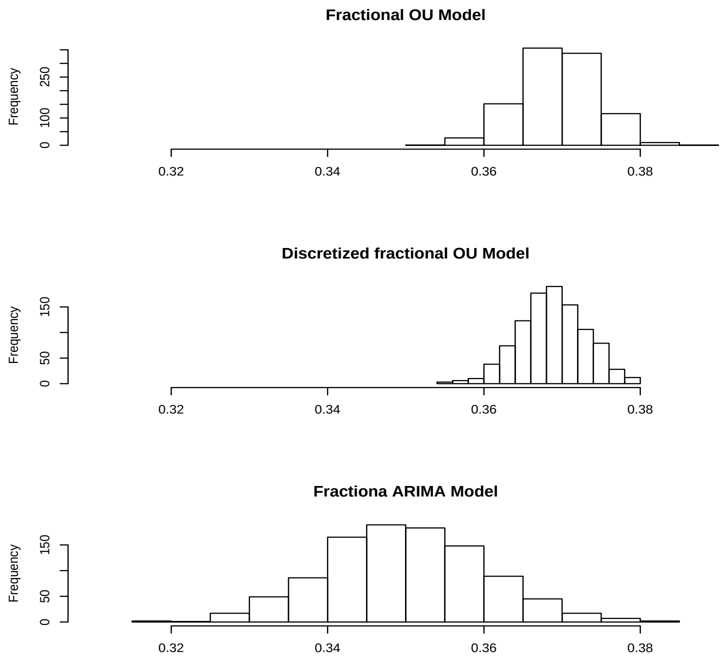


Figure 6: Empirical distribution of the unobserved volatility using the three models, when the "true" underlying data are described by a fractional OU model.

Strike Price	fractional OU	Discretized fOU	fractional ARIMA
90	72.187	72.203	72.185
100	62.335	62.331	62.253
110	52.707	52.779	52.540
120	43.665	43.594	43.750
130	34.863	34.946	34.759
140	27.785	27.480	26.663
150	21.471	21.407	20.645
160	15.307	15.6914	15.276
170	11.730	11.577	10.148
180	8.4018	8.3831	8.0326
190	5.8517	5.6555	5.3244
200	4.1847	4.1151	3.5851
210	2.8696	2.6276	2.0635

Table 8: Call Option Prices for a simulated model using the three models, when the "true" underlying data are described by a fractional OU model.

Model	**Implied $\hat{\mathbf{H}}$**
Continuous fOU	0.53
Discretized fOU	0.53
fARIMA	0.52

Table 9: Implied H for the S&P 500 on March 30^{th} 2009, under the three models.

Strike Price	Continuous fOU	Discretized fOU	fractional ARIMA
670	123.96	125.66	129.54
680	115.82	117.54	119.76
690	107.94	108.53	111.79
700	100.32	101.34	104.40
710	92.99	94.05	97.36
720	85.98	86.59	89.52
730	79.26	80.87	83.31
740	72.87	75.17	76.28
750	66.81	69.18	71.23
760	61.07	63.67	65.41
770	55.67	58.00	60.11
780	50.59	52.35	55.03
790	45.85	47.69	51.23
800	41.43	42.69	46.09
810	37.33	39.13	41.63
820	33.54	35.09	37.79
830	30.05	31.24	34.31
840	26.83	27.28	30.70
850	23.90	24.85	27.88

Table 10: Call Option Prices on the S&P 500 with $S_0 = \$787.5$, $r = 0.21\%$ and $T = 35$ business days.

Now, using the implied values of H and the corresponding estimated parameters for each model, we price the call option announced above. The option prices are presented in Table 10 and a graphical comparison with the continuous fractional OU model can be seen in Figure 7.

We observe that both discrete time models are close to the continuous time one. Although we computed a different implied H for the fractional ARIMA model, in Figure 7 it seems to be slightly closer to the continuous time fractional OU model. However, if we compare the mean square error of the option prices with the center of the bid-ask spread for the three models, then the continuous time model is significantly better than the other two (has a lower MSE), the second best being the discretized fractional OU model and the last is the fractional ARIMA model. This is empirical evidence that the continuous-time fractional OU

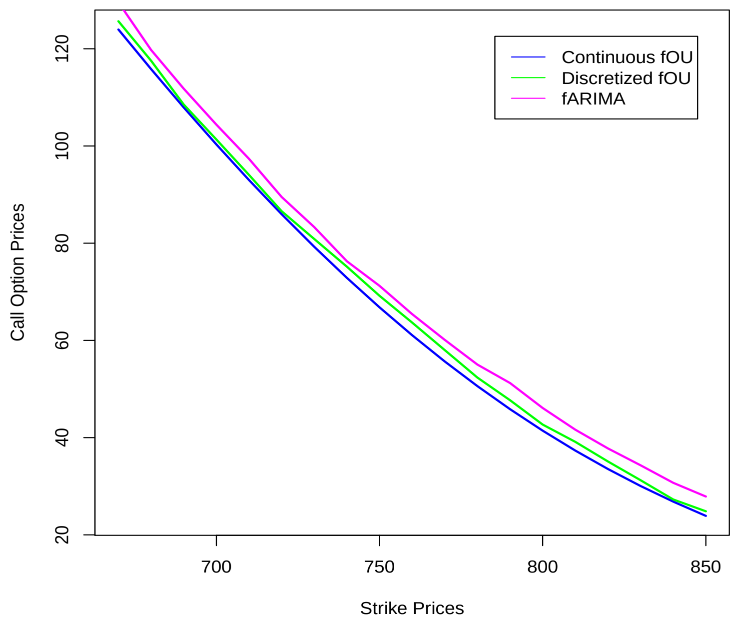


Figure 7: Comparison of Call Option Prices on the S&P 500 using the three models.

model is superior in option-pricing modeling terms irrespective of computational time, and also confirms the conclusions of the mis-specification study.

5 Conclusions and recommendations

This article studies statistical memory in financial volatility, and its effect on option prices in liquid markets, where high-frequency data is available. It proposes several long-memory stochastic models for stock and index prices, and implements a calibration framework to determine the memory length (value of the long-memory parameter). The main calibration criterion is the consistency of calculated call option prices with those observed on the option market.

In accordance with established practice in continuous-time finance, martingale extensions of the Black-Scholes framework are used: the basic continuous-time model is that of Comte and Renault [14], in which the volatility is an autonomous stochastic process, namely a fractional Ornstein-Uhlenbeck (OU) process, which has the desirable properties of being mean-reverting and having long memory. The resulting model is arbitrage-free. This article also proposes two discrete models for the purpose of reducing the computational time for parameter estimation and option pricing. One is a discretization of the Comte-Renault model (discretized fractional OU model), the other is a fractional ARIMA model proposed by Harvey [31] and Breidt et al. [9].

Parameter estimation and option pricing should be based on the premise that option market makers are forward-looking, in their expectations of where the market volatility might be heading. This warrants this article's integrated technique which trusts current option data to determine whether the stock market's volatility has any statistical memory, and which uses the option data to determine simultaneously the volatility memory parameter and the current option prices based on this parameter. The other model parameters are estimated using standard statistical methodology, for each possible value of the long-memory parameter, whose final value is determined by minimizing the distance between computed option prices and observed market prices. Several minimizing schemes can be used, depending on observed trade volumes or on user preferences. The option-pricing step is based on the multinomial recombining tree proposed by Florescu and Viens [23] for generic random volatility settings, combined with the long-memory stochastic volatility filter implemented in [13].

By working with both simulated data and call option data from the S&P 500 index, the article determines that the continuous model is the most accurate one. The best estimation and pricing results, as judged by comparisons to market prices, are obtained by using trade volumes to weigh these prices. Local minimizing schemes based on restricted ranges of option strike prices also produce good results, even when volatility is tracked using a moderate frequency of observations.

More generally, the empirical volatility distribution obtained from the stochastic volatility filter, and as a consequence, the calibrated value of the long memory parameter, are rather insensitive to the frequency of observations. However, the entire integrated estimation and pricing technique requires high-frequency data in order to estimate the non-memory model parameters.

In order to have manageable CPU times, discrete models must be used. The discretized fractional OU model performs better than the fractional ARIMA model at all observation frequencies. This is no surprise, since the best fit of all is obtained from the more computationally intensive continuous-time fractional OU model, which therefore provides the best description of the underlying phenomenon. In conclusion, in order to have reasonable CPU times and accurate option prices, in markets where the volatility is suspected to have statistical long memory, the discretized fractional OU model is recommended.

References

[1] Amin, K.I. and Ng, V. (1993). Option valuation with systematic stochastic volatility. *Journal of Finance*, **48**:881-910.

[2] Andersen, T. and Bollerslev, T. (1997). Heterogeneous information arrivals and return volatility dynamic: Incovering the long-run in high-frequency returns. *Journal of Finance.*, **52**, 975–1005.

[3] Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996). Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity. *J. Econometrics* **74**(1):3–30.

[4] Beran, J. (1994): *Statistics for Long-Memory Processes.* Chapman and Hall.

[5] Black, F. and Scholes, M. (1973): The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* **81** (3): 637—654.

[6] Bollerslev, T. (1986). A generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**:307-327.

[7] Bollerslev T., Chou, R. and Kroner, K. (1992). ARCH modelling in finance. *Journal of financial econometrics.* **52**(5).

[8] Bollerslev, T. and Mikkelsen, H. (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, **73**, 151–184.

[9] Breidt, F. J., Crato, N. and De Lima, P. (1998). The Detection and Estimation of Long-Memory in Stochastic Volatility. *Journal of Econometrics.* **83**:325-348.

[10] Casas, I. and Gao, J. (2008) Econometric estimation in long-range dependent volatility models: Theory and practice. *Journal of Econometrics* **147**:72-83.

[11] Cheridito, P., Kawaguchi, H. and Maejima, M. (2003). Fractional Ornstein-Uhlenbeck processes. *Electronic Journal of Probability.* **8**(3):1-14.

[12] Chronopoulou, A. (2011). Parameter estimation and calibration for long-memory stochastic volatility models. *Handbook of Modeling High-Frequency Data in Finance*. To appear.

[13] Chronopoulou, A. and Viens, F. (2010). Estimation and pricing under long-memory stochastic volatility. *Annals of Finance*.

[14] Comte, F. and Renault, E. (1998). Long Memory in Continuous-time Stochastic Volatility Models. *Mathematical Finance*, **8**(4):291-323.

[15] Comte, F., Coutin, L., and Renault, E. (2010). Affine Fractional Stochastic Volatility Models. To appear in *Annals of Finance*.

[16] Cox, J., Ingersoll, J. Jr. and Ross, S. (1985). An intertemporal general equilibrium model of asset prices. *Econometrica*, **53**:363-384.

[17] De Lima, P. and Crato, N. (1994). Long-range dependence in the conditional variance of stock returns. *Economics Letters*, **45**(3), 281–285 .

[18] Del Moral, P., Jacod, J. and Protter, P. (2001).The Monte-Carlo method for filtering with discrete time observations. *Probability Theory and Related Fields*, **120**:346-368.

[19] Deo, R.S. and Hurvich, C.M. (2001). On the log periodogram regression estimator of the memory parameter in long memory stochastic volatility models. *Econometric Theory*, **17**, 686–710.

[20] Ding, Z. C., Granger, W., J. and Engle, R. F. (1993). A Long Memory Property of Stock Market Returns and a New Model. *J. Empirical Finance* **1**, 1.

[21] Duan, J.C. (1995). The garch option pricing model. *Mathematical finance*, **5**:13-32.

[22] Engle, R.F. and Mustafa, C. (1992). Implied ARCH models for option prices. *Journal of Econometrics*, **52**, 289-311.

[23] Florescu, I. and Viens, F. G. (2008).Stochastic volatility: option pricing using a multinomial recombining tree. *Applied Mathematical Finance*, **15**(2), 151–181.

[24] Fouque, J.-P., Papanicolaou, G. and Sircar, K. R. (2000a).*Derivatives in Financial Markets with Stochastic Volatility.* Cambridge University Press.

[25] Fouque, J.-P., Papanicolaou, G. and Sircar, K. R. (2000b). Mean-reverting stochastic volatility. *International Journal of Theoretical and Applied Finance,* **3**(1):101-142.

[26] Gao, J., Ahn, V. and Heydy, C. (2002). Statistical estimation of nonstationary Gaussian processes with long-range dependence and intermittency. *Stochastic Processes and Applications. 99*, 295–321.

[27] Gao, J., Ahn, V., Heydy, C. and Tieng, Q. (2001). Parameter estimation fo stochastic processes with long-range dependence and intermittency. *Journal of Time Series Analysis. 4*, 221–238.

[28] Geweke, J. and S. Porter-Hudak (1983).The estimation and application of long-memory time-series models. *Journal of timeseries analysis*, **4**, 221-237.

[29] Granger, C.W., Joyeux, R., 1980. An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis* **1**, 15–29.

[30] Hamilton, J.D. (1994) *Time Series Analysis.* Princeton University Press.

[31] Harvey, A.C. (1998). Long Memory in Stochastic Volatility, in J.Knight and S. Satchell (eds.) *Forecasting Volatility in Financial Markets*, Butterworth-Haineman, Oxford, pp. 307-320.

[32] Henry, M. and Payne, R. (1997). An investigation of long-range dependence in intra-day foreign exchange rate volatility. Discussion paper 264,*Financial Markets Group, London School of Economics.*

[33] Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatility. *Journal of Finance,* **42**:281-300.

[34] Hurvich, C. M. and Soulier Ph. (2009). Volatility models with long memory. In T. Andersen, R. Davis, J.-P. Kneiss and Th. Mikolsch (Eds.) *Handbook of financial time series.* Springer.

[35] Hurst, H.E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, **116**: 770-799.

[36] Lobato, I. and Savin, N. E. (1998). Real and Spurious Long-Memory Properties of Stock-Market Data. *Journal of Business and Economic Statistics.* **16**:261-283.

[37] Merton, R. C. (1973). Theory of Rational Option Pricing. *Bell Journal of Economics and Management Science (The RAND Corporation)* **4** (1): 141–183.

[38] Neuenkirch, A. (2006). Optimal approximation of SDE's with additive fractional noise. *Journal of Complexity*, **22**(4):459-474.

[39] Nualart, D. *Malliavin Calculus and Related Topics*, 2nd edition, Springer Verlag, 2006.

[40] Taylor, S.J. (1982) Financial returns modelled by the product of two stochastic processes, as study of daily sugar prices 1961-79. in *O.D. Anderson (ed.) Timeseries Analysis: Theory andd Practice 1*, North Holland, Amsterdam, 203–226.

[41] Taylor, S.J. (1986). *Modeling financial time series.* John Wiley, Chichester, UK.

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