Online Particle Filtering of Stochastic Volatility

Hana Baili

Online Particle Filtering of Stochastic Volatility

Hana Baili *

Abstract—**A method for online estimation of the volatility when observing a stock price is proposed. This is based on modeling the volatility dynamics as a stochastic differential equation that is constructed using a technique from the control theory [1]. Identification of the model parameters using the observations is proposed afterwards [2]. It is based on some stochastic calculus. Volatility estimation is then reformulated as a filtering problem. An alternative filter instead of the optimal one is proposed since the latter is not computationally feasible. It is based on samples (or particles) drawn by discretization of the stochastic volatility model. Besides, the main feature that makes online particle filtering possible is analytic resolution of the Fokker-Planck equation for the current return. To the best of our knowledge, such technique for modeling together with online filtering of the volatility are quiet novel. The method is implemented on real data: the Heng Seng index price; this shows a period of relatively high volatility that corresponds obviously to the Asiatic crisis of October 1997.**

Keywords: stochastic volatility, stochastic differential equations, Fokker-Planck equation, particle filtering.

1 Introduction

Let $S = (S_t)_{t\in\mathbb{R}_+}$ be an $\mathbb{R}_+$-valued semimartingale based on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P})$ which is assumed to be continuous. The process S is interpreted to model the price of a stock. A basic problem arising in Mathematical Finance is to estimate the price volatility, i.e. the square of the parameter σ in the following stochastic differential equation

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

where $W = (W_t)_{t\in\mathbb{R}_+}$ is a Wiener process. It turns out that the assumption of a constant volatility does not hold in practice. Even to the most casual observer of the market, it should be clear that volatility is a random function of time which we denote σ_t^2. Itô's formula for the return $y_t = \log(S_t/S_0)$ yields

$$dy_t = \left(\mu - \frac{\sigma_t^2}{2}\right) dt + \sigma_t\, dW_t \quad y_0 = 0 \tag{1}$$

The main objective is to estimate in discrete real-time one and only one particular sample path of the volatility process using one and only one observed sample path of the return. As regards the drift μ, it is constant but unknown. Under the so-called risk-neutral measure, the drift is a riskless rate which is well known; actually one finds that μ does not cancel out, for instance, when calculating conditional expectations in a filtering problem. For this argument no change of measure is required, we work directly in the original measure $\mathbb{P}$, and μ has to be estimated from the observed sample path of the return as well.

*École Supérieure d'Électricité, Department of Signal Processing and Electronic Systems, 3 rue Joliot Curie, Plateau de Moulon, 91192 Gif sur Yvette, France. Tel/Fax: 0033 1 6985 1428/1429 Email: hana.baili@supelec.fr

2 A model for the stochastic volatility

We assume prior information about the unknown process σ_t^2 of instantaneous volatility: a parametric model for its autocorrelation function

$$\gamma(\tau) = D\exp(-\alpha|\tau|) \quad \tau \in \mathbb{R} \tag{2}$$

for some $\alpha > 0$. This type of autocorrelation function includes short-term or middle-term memory in the correlation pattern of the volatility. Then the spectral density of σ_t^2 is given by the formula

$$\Gamma(\omega) = \frac{1}{2\pi}\int_{\mathbb{R}} \gamma(\tau)\exp(-j\omega\tau)d\tau = \frac{1}{2\pi}\frac{2D\alpha}{\omega^2 + \alpha^2}$$

where $j = \sqrt{-1}$. The spectral density $\Gamma(\omega)$ is rewritten as

$$\Gamma(\omega) = \frac{1}{2\pi}\left|\frac{H(j\omega)}{F(j\omega)}\right|^2 \quad \omega \in \mathbb{R}$$

where $H(j\omega) = \sqrt{2D\alpha}$ and $F(j\omega) = j\omega + \alpha$. Notice now that

$$\Phi(s) = \frac{H(s)}{F(s)} \quad s \in C$$

represents the transfer function of some stationary linear system; this system is furthermore stable as the root of $F(s)$ is in the left half-plane of the complex variable s. Recalling that $1/2\pi$ is the spectral density of a white noise of intensity 1, we come to the conclusion that σ_t^2 may be considered as the response of the filter whose transfer function is $\Phi(s)$, to a white noise with unit intensity. The differential equation describing such a filter is

$$\dot{u}(t) + \alpha u(t) = \sqrt{2D\alpha}\, w(t)$$

where $w(t)$ and $u(t)$ are respectively the input and the output of the filter. Set $x_t = u(t)$, then the process σ_t^2—denoted x_t in the following—solves the SDE

$$dx_t = -\alpha x_t\, dt + \sqrt{2D\alpha}\, d\tilde{W}_t \tag{3}$$

with reflection at 0 so as to assure the positivity; $\tilde{W} = (\tilde{W}_t)_{t\in\mathbb{R}_+}$ is a Wiener process, and W and $\tilde{W}$ are independent. We shall freely call (3) our stochastic volatility model.

3 Filtering

Now we consider the filtering problem associated to the couple (x_t, y_t): we have noisy nonlinear observations of x_t, the $\mathbb{R}$-valued discrete-time process of returns $(y_n)_{n=1,2,...}$ indexed at irregularly spaced instants $t_1, t_2, ...$ The observation times are assumed to be rigourously determined. The observations process is related to the state process $(x_t)_{t\in\mathbb{R}_+}$ via the conditional distribution

$$\mathbb{P}\{y_n \in \Gamma | y_1, ..., y_{n-1}, (x_t : 0 \le t \le t_n)\} \quad n \ge 1$$

for Γ a Borel-measurable set from $\mathbb{R}$. For homogeneity of notation we set $t_0 = 0$ so that $y_{n=0} = y_{t=t_0} = 0$. Now look at the distribution above and recall that $y_n = y(t_n)$ and that the process y_t solves the SDE

$$dy_t = \left(\mu - \frac{x_t}{2}\right) dt + \sqrt{x_t}\, dW_t \quad y_0 = 0 \tag{4}$$

This is (1) where σ_t is denoted $\sqrt{x_t}$. For $t \ge t_{n-1}$

$$y_t = y_{n-1} + \int_{t_{n-1}}^{t} \left(\mu - \frac{x_s}{2}\right) ds + \int_{t_{n-1}}^{t} \sqrt{x_s}\, dW_s \tag{5}$$

and thus

$$\mathbb{P}\{y_n \in \Gamma | y_1, ..., y_{n-1}, (x_t : 0 \le t \le t_n)\} = \mathbb{P}\{y_n \in \Gamma | y_{n-1}, (x_t : t_{n-1} \le t \le t_n)\}$$

Given a sample path of $(x_t)_{t_{n-1}\le t\le t_n}$ and the observation y_{n-1}, $(y_t)_{t_{n-1}\le t\le t_n}$ is a Markov process with state space $\mathbb{R}$ satisfying (5). This leads to the central concept of this section: the Fokker-Planck equation [3]. The domain of the Fokker-Planck operator for the return,

$$\mathbf{FP}p(t,y) = \left(\frac{x_t}{2} - \mu\right)\frac{\partial p}{\partial y}(t,y) + \frac{x_t}{2}\frac{\partial^2 p}{\partial y^2}(t,y),$$

is the set of probability density functions $p(t,y)$ satisfying: for any $t \ge 0$ such that $x_t \ne 0$, $p(t,y)$ must be null as $y \to \pm\infty$. This is absorption on the boundary of $\mathbb{R}$, the state space of the return. Then, given a sample path of $(x_t)_{t_{n-1}\le t\le t_n}$ and the observation y_{n-1}, the distribution density of y_t solves the Fokker-Planck equation

$$\frac{\partial p}{\partial t}(t,y) = \mathbf{FP}p(t,y) \quad t_{n-1} < t \le t_n \tag{6}$$

with the initial condition $p(t_{n-1}, y) = \delta(y - y_{n-1})$. The formal solution of the above partial differential equation is

$$p(t,y) = \exp\{(t - t_{n-1})\mathbf{FP}\}\, p(t_{n-1}, y)$$

Since $\mathbf{FP}$ is a sum of two non commuting operators, the exponential operator $\exp\{(t - t_{n-1})\mathbf{FP}\}$ cannot be expressed as simple products of terms involving each of these. Nevertheless, the solution of the Fokker-Planck equation is obtained using the Trotter product formula [4]. For two arbitrary operators A and B

$$\exp\{t(A+B)\} = \lim_{n\to\infty}\left(\exp\left\{\frac{t}{n}A\right\}\exp\left\{\frac{t}{n}B\right\}\right)^n$$

Then the solution of (6) is

$$p(t,y) = \lim_{n\to\infty}\left(\exp\left\{\frac{\rho(t-t_{n-1})}{n}\frac{d}{dy}\right\}\exp\left\{\frac{\epsilon(t-t_{n-1})}{n}\frac{d^2}{dy^2}\right\}\right)^n \delta(y - y_{n-1})$$

where

$$\rho = \frac{x_t}{2} - \mu \quad \epsilon = \frac{x_t}{2}$$

For algebraic manipulations we use the integral representation of the delta function and write the solution of (6) as

$$p(t,y) = \lim_{n\to\infty}\Theta^n \frac{1}{2\pi}\int_{z\in\mathbb{R}} \exp\{-jzy\}\exp\{jzy_{n-1}\}$$

where

$$\Theta = \exp\left\{\frac{\rho(t-t_{n-1})}{n}\frac{d}{dy}\right\}\exp\left\{\frac{\epsilon(t-t_{n-1})}{n}\frac{d^2}{dy^2}\right\}$$

We claim that

$$\exp\left\{\frac{\epsilon(t-t_{n-1})}{n}\frac{d^2}{dy^2}\right\}\exp\{-jzy\} = \exp\left\{-\frac{\epsilon(t-t_{n-1})}{n}z^2 - jzy\right\}$$

$$\exp\left\{\frac{\rho(t-t_{n-1})}{n}\frac{d}{dy}\right\}\exp\{-jzy\} = \exp\left\{-\frac{\rho(t-t_{n-1})}{n}jz - jzy\right\}$$

Therefore

$$\Theta\exp\{-jzy\} = \exp\left\{-\frac{\epsilon(t-t_{n-1})}{n}z^2 - \frac{\rho(t-t_{n-1})}{n}jz - jzy\right\}$$

$$\Theta^n\exp\{-jzy\} = \exp\left\{-\epsilon(t-t_{n-1})z^2 - \rho(t-t_{n-1})jz - jzy\right\}$$

and thus

$$p(t,y) = \frac{1}{2\pi}\int_{z\in\mathbb{R}} \exp\left\{-\epsilon(t-t_{n-1})z^2 + jz\left[-y + y_{n-1} - \rho(t-t_{n-1})\right]\right\}$$

Let X be a Gaussian random variable with mean $\bar{x}$ and variance $\bar{v}$. Let $\psi(u)$, $u \in \mathbb{R}$, be its characteristic function, i.e.

$$\begin{aligned}\psi(u) &= \mathbb{E}[\exp\{juX\}] \\ &= \frac{1}{\sqrt{2\pi\bar{v}}}\int_{-\infty}^{+\infty}\exp\{juz\}\exp\left\{-\frac{(z-\bar{x})^2}{2\bar{v}}\right\}dz \\ &= \exp\left\{j\bar{x}u - \frac{\bar{v}\,u^2}{2}\right\}\end{aligned}$$

Then

$$p(t,y) = \frac{1}{2\sqrt{\pi\epsilon(t-t_{n-1})}}\,\psi\left(-y+y_{n-1}-\rho(t-t_{n-1})\right)$$

with

$$\bar{x}=0 \quad \bar{v}=\frac{1}{2\epsilon(t-t_{n-1})}$$

and hence we obtain for $t_{n-1} \le t \le t_n$

$$p(t,y) = \frac{1}{\sqrt{2\pi x_t(t-t_{n-1})}} \times \exp\left\{-\frac{\left[-y+y_{n-1}+\left(\mu-\frac{x_t}{2}\right)(t-t_{n-1})\right]^2}{2x_t(t-t_{n-1})}\right\}$$

3.1 Conditional density characterization: the optimal filter

The optimal estimate—in a sense of the mean square—of $f(x_t)$ given the observations $y_1,...,y_{n-1}$ up to time t is the conditional expectation

$$\mathbb{E}\left[f(x_t)|y_1,...,y_{n-1}\right] \quad t_{n-1} \le t < t_n \quad n \ge 1$$

for all reasonable functions f on $\mathbb{R}_+$. We assume that $\mathbb{P}\{x_t \le x|y_1,...,y_{n-1}\}$ possesses a density with respect to the Lebesgue measure λ on $\mathbb{R}_+$:

$$\Pi_{x_t|y_1,...,y_{n-1}}(x) = \frac{d\mathbb{P}\{x_t \le x|y_1,...,y_{n-1}\}}{\lambda(dx)}$$

Now look at the SDE (3), the Fokker-Planck operator for x_t is

$$\mathbf{FP}p(x) = \alpha p(x) + \alpha x p'(x) + D\alpha p''(x)$$

The domain of this operator is the set of distribution densities $p(x)$ on $\mathbb{R}_+$ satisfying

$$p'(0) = 0$$

This is due to the reflection of the process x_t on the boundary $\{0\}$ of its phase space $\mathbb{R}_+$.

It follows that the posterior distribution density $\Pi_{x_t|y_1,...,y_{n-1}}(t,x)$ for $t_{n-1} \le t < t_n$, $n \ge 1$, solves the Fokker-Planck equation

$$\frac{\partial p}{\partial t}(t,x) = \mathbf{FP}p(t,x) \quad t_{n-1} < t < t_n$$

i.e.

$$\frac{\partial p}{\partial t}(t,x) = \alpha p(t,x) + \alpha x\frac{\partial p}{\partial x}(t,x) + D\alpha\frac{\partial^2 p}{\partial x^2}(t,x) \quad t_{n-1} < t < t_n \tag{7}$$

with the initial condition

$$p(t_{n-1},x) = \Pi_{x(t_{n-1})|y_1,...,y_{n-1}}(x) \tag{8}$$

and the boundary condition

$$\frac{\partial p}{\partial x}(t,0) = 0 \tag{9}$$

This is a static relation for $x = 0$, i.e., it holds for any $t \in [t_{n-1}, t_n[$.

At each observation instant t_n, $n \ge 1$, $\Pi_{x(t_n)|y_1,...,y_n}(x)$ solves the Bayes rule

$$\Pi_{x(t_n)|y_1,...,y_n}(x) \propto \Pi_{x(t_n^-)|y_1,...,y_{n-1}}(x)\Pi_{y_n|y_1,...,y_{n-1},x(t_n)=x}(y_n) \tag{10}$$

where

$$\Pi_{y_n|y_1,...,y_{n-1},x(t_n)=x}(y_n) = \frac{1}{\sqrt{2\pi x(t_n-t_{n-1})}} \times \exp\left\{-\frac{\left[-y_n+y_{n-1}+(\mu-\frac{x}{2})(t_n-t_{n-1})\right]^2}{2x(t_n-t_{n-1})}\right\}$$

and $\Pi_{x(t_n^-)|y_1,...,y_{n-1}}(x)$ is the solution of (7-9) as $t \uparrow t_n$.

4 Identification

It follows from (4) that the variation process $[y]_t$ of y_t is given by

$$[y]_t = \int_0^t x_s\,ds$$

thus

$$[y]_{t_n} - [y]_{t_{n-1}} = \int_{t_{n-1}}^{t_n} x_s\,ds \quad n = 1,2,...$$

On the other hand, so long as every duration between two successive observations is small, the following approximation holds

$$[y]_{t_n} \approx \sum_{i=1}^{n}(y_i - y_{i-1})^2$$

Thus

$$\int_{t_{n-1}}^{t_n} x_s\,ds \approx (y_n - y_{n-1})^2$$

i.e., the couple of series below coincide approximatively

$$S = \left\{\int_{t_{n-1}}^{t_n} x_s\,ds\right\}_{n=1,2,...} \qquad S' = \left\{(y_n - y_{n-1})^2\right\}_{n=1,2,...}$$

and so do their autocorrelation functions. The following is the computation of the autocorrelation function for S, i.e., the series of aggregations of the instantaneous volatility on the observation intervals. To do this we need to have $t_n - t_{n-1} = \delta$ for each $n = 1, 2, ...$ and as mentioned above δ must be small (we set $\delta = 1$ time unit). Then for $k = 1, 2, ...$

$$\mathbb{E}\left[\int_{t_{n-1}}^{t_n} x_u\, du \times \int_{t_{n-k-1}}^{t_{n-k}} x_v\, dv\right] = \int_{t_{n-1}}^{t_n} \int_{t_{n-k-1}}^{t_{n-k}} \gamma(u-v)\, du\, dv$$

If we replace γ by its expression in (2), we obtain the following formula for $k = 1, 2, ...$

$$\mathbb{E}\left[\int_{t_{n-1}}^{t_n} x_u\, du \times \int_{t_{n-k-1}}^{t_{n-k}} x_v\, dv\right] = \quad (11)$$
$$\frac{D}{\alpha^2}\left(\exp\{-\alpha\delta(k-1)\} - 2\exp\{-\alpha\delta k\} + \exp\{-\alpha\delta(k+1)\}\right)$$

It follows that D and α may be obtained by least squares of the difference between the autocorrelation function of S', calculated from the observations, and the autocorrelation function given by formula (11).

The following gives an approximation for the drift parameter μ in (1).

$$y_n - y_{n-1} = \int_{t_{n-1}}^{t_n} \left(\mu - \frac{x_s}{2}\right) ds + \int_{t_{n-1}}^{t_n} \sqrt{x_s}\, dW_s$$

implies that

$$\mathbb{E}\left[y_n - y_{n-1}\right] = \mu\,\delta - \frac{1}{2}\,\mathbb{E}\left[\int_{t_{n-1}}^{t_n} x_s\, ds\right]$$

But

$$\int_{t_{n-1}}^{t_n} x_s\, ds \approx (y_n - y_{n-1})^2$$

thus

$$\mathbb{E}\left[y_n - y_{n-1}\right] \approx \mu\,\delta - \frac{1}{2}\,\mathbb{E}\left[(y_n - y_{n-1})^2\right]$$

i.e.

$$\mu \approx \frac{1}{\delta}\left(\mathbb{E}\left[y_n - y_{n-1}\right] + \frac{1}{2}\,\mathbb{E}\left[(y_n - y_{n-1})^2\right]\right) \quad (12)$$

The daily price of the Hang Seng index of the market of Hong Kong is observed during 3191 successive trading days from 1995 to 2007. This is plotted in Figure 1. Figure 2 shows the daily returns

$$y_n - y_{n-1} = \log\left(\frac{S_{t_n}}{S_{t_{n-1}}}\right) \quad n = 1, ..., 3190$$

For the approximation of the drift μ in (12) we get $5.4008e-004$. The second order moment D and the rate α that give a good fitting between the autocorrelation function of S and its empirical approximation are $3.5926e-007$ and 0.0857 respectively. The SDE (3) for the stochastic volatility of the stock is thus calibrated, and we now go back to filtering.

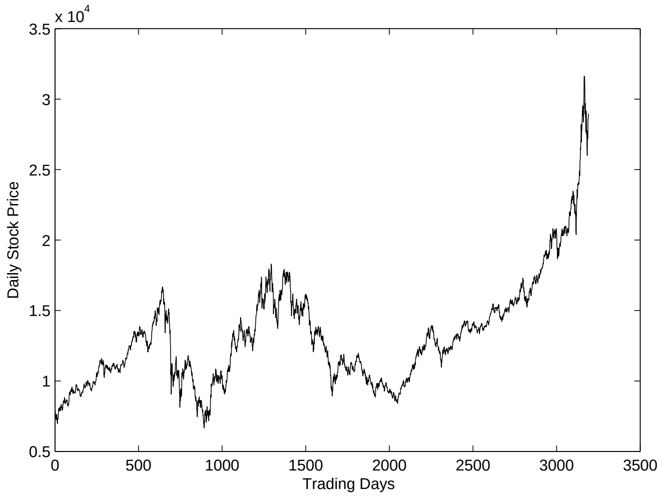

Figure 1:

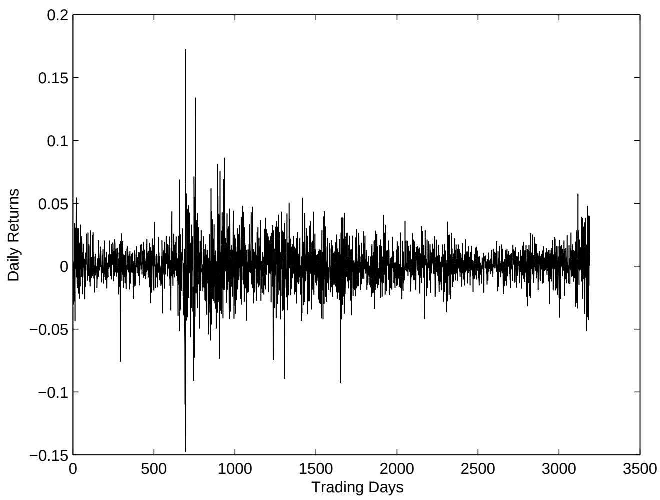

Figure 2:

5 A Monte-Carlo particle filter

The true filter (7-10) which is optimal in a mean square sense involves a resolution of the Fokker-Planck equation. Both analytic and numerical solutions for this partial differential equation are computationally intractable. This drives us to an alternative Monte-Carlo filter [5]. We wish to approximate the posterior distribution as a weighted sum of random Dirac measures: for Γ a Borel-measurable set from $\mathbb{R}_+$

$$\mathbb{P}\{x_t \in \Gamma | y_1, ..., y_{n-1}\} \approx \sum_{k=1}^{K} w_k \, \epsilon_{\xi_k}(\Gamma) \quad t_{n-1} \leq t < t_n \quad n \geq 1$$

where the particles ξ_k are independent identically distributed random variables with "the same" law as x_t; these particles are indeed samples drawn from the Euler discretization of the SDE (3). Here we use the well known Euler scheme since there isn't a significant gain with more sophisticated discretization schemes.
Then, for any function f on $\mathbb{R}_+$

$$\mathbb{E}\left[f(x_t) | y_1, ..., y_{n-1}\right] \approx \sum_{k=1}^{K} w_k f(\xi_k) \quad t_{n-1} \leq t < t_n \quad n \geq 1$$

The weights $\{w_k\}_{k=1,...,K}$ are updated only as and when an observation y_n proceeds, each one according to the likelihood of its corresponding particle, i.e., at each observation time t_n

$$w_k = \frac{\Pi_{y_n | y_1, ..., y_{n-1}, x(t_n)=\xi_k}(y_n)}{\sum_{\ell=1}^{K} \Pi_{y_n | y_1, ..., y_{n-1}, x(t_n)=\xi_\ell}(y_n)}$$

where $\{\xi_k\}_{k=1,...,K}$ are samples with the same law as $x(t_n)$.

Besides sampling, there may be (importance) resampling at each observation time: the set of particles is updated for removing particles with small weights and duplicating those with important weights. We simulate K new iid random variables according to the distribution

$$\sum_{k=1}^{K} w_k \, \epsilon_{\xi_k}$$

Obviously, the new particles have new weights and thus give a new approximation for the posterior distribution. On the other hand, these new particles are used to initialize the Euler discretization scheme for the next sampling.

The following is the remainder of implementation details of the Monte-Carlo particle filter.

- Number of particles: $K = 1000$
- Time step of the Euler discretization: 0.01 time unit
- Since the distribution of the initial volatility x_0 is not available in practice, let us take a uniform distribution on $[0, 1]$; its density satisfies the imposed condition (9).

The sample path of the square root volatility (in percent) of the Heng Seng index price is displayed in Figure 3. This sample path exhibits relatively high volatilities that are clustered together round the 697th trading day; this corresponds to the Asian financial crisis of October 1997.

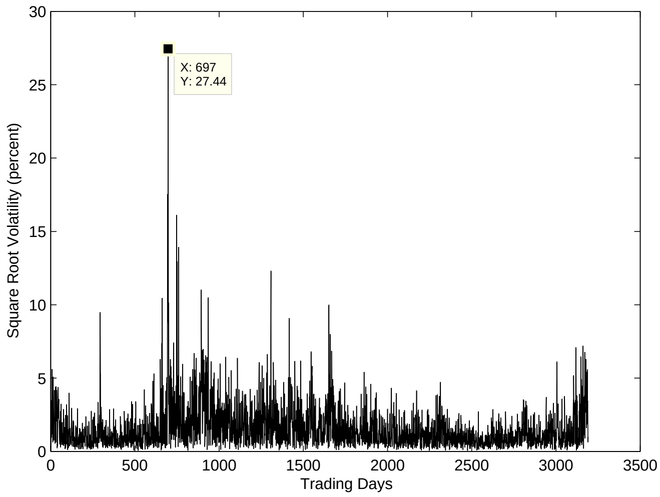

Figure 3:

6 Conclusion

Probabilistic management of uncertainty in dynamical systems is proposed when illustrated on an application from financial engineering: volatility estimation. We consider the volatility as a stochastic process and construct a filter that is recursive and pathwise in the observations; these two aspects are designated by the term online (or real-time) filtering. The filter output is thus one—and only one—particular sample path of the volatility process. Besides, the main feature that makes online particle filtering possible is analytic resolution of a Fokker-Planck equation. It is worth noting that our method does not need any effort to transform data, for example, to take off seasonality. The conformity between the implementation result—within a low simulation cost—and some practical issues prove to my satisfaction the performance of the method.

Acknowledgement. I am grateful to Professor Sana Louhichi for her stimulating interest and many discussions motivating the research.

References

[1] Pugachev, V.S., Sinitsyn, I.N., *Stochastic Systems, Theory and Applications*, John Wiley & Sons, 1987.

[2] Baili, H., "Uncertainty management for estimation in dynamical systems," *IEEE Asia Pacific Conference on Circuits and Systems*, Singapore, pp. 1992-1995, 12/06

[3] Gardiner, C.W., *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, Second Edition, Springer-Verlag, 1985.

[4] Valsakumar, M.C., "Solution of Fokker-Planck equation using Trotter's formula," *Journal of Satistical Physics*, V32, N3, pp. 545-553, 09/83

[5] Baili, H., Snoussi, H., "Stochastic Filetring with Networked Sensing," *IEEE 70th Vehicular Technology Conference*, Anchorage, Alaska, CD-ROM Proceedings, 09/09