Filtering in Finance

December 20, 2002

Alireza Javaheri[1], Delphine Lautier [2], Alain Galli[3]

Abstract

In this article we present an introduction to various Filtering algorithms and some of their applications to the world of Quantitative Finance. We shall first mention the fundamental case of Gaussian noises where we obtain the well-known Kalman Filter. Because of common nonlinearities, we will be discussing the Extended Kalman Filter (EKF) as well as the Unscented Kalman Filter (UKF) similar to Kushner's Nonlinear Filter. We also tackle the subject of Non-Gaussian filters and describe the Particle Filtering (PF) algorithm. Lastly, we will apply the filters to the term structure model of commodity prices and the stochastic volatility model.

1 Filtering

The concept of filtering has long been used in Control Engineering and Signal Processing. Filtering is an *iterative* process that enables us to estimate a model's parameters when the latter relies upon a large quantity of observable *and* unobservable data. The Kalman Filter is fast and easy to implement, despite the length and noisiness of the input data.
We suppose we have a temporal time-series of observable data $\mathbf{z}_k$ (*e.g.* stock prices [17], [31], interest rates [5], [26], futures prices [21], [22]) and a model using some unobservable time-series $\mathbf{x}_k$ (*e.g.* volatility, correlation, convenience yield) where the index k corresponds to the time-step. This will allow us to construct an algorithm containing a transition equation linking two consecutive unobservable states, and a measurement equation relating the observed data to this hidden state.
The idea is to proceed in two steps: first we estimate the hidden state *a priori* by using all the information prior to that time-step. Then using this predicted value together with the new observation, we obtain a conditional *a posteriori* estimation of the state.

In what follows we shall first tackle linear and nonlinear equations with Gaussian noises. We then will extend this idea to the Non-Gaussian case.

[1]RBC Capital Markets
[2]Université Paris IX, Ecole Nationale Supérieure des Mines de Paris
[3]Ecole Nationale Supérieure des Mines de Paris

Further, we shall provide a mean to estimate the model parameters via the maximization of the *likelihood* function.

1.1 The Simple and Extended Kalman Filters

1.1.1 Background and Notations

In this section we describe both the traditional Kalman Filter used for linear systems and its extension to nonlinear systems known as the Extended Kalman Filter (EKF). The latter is based upon a first order linearization of the transition and measurement equations and therefore would coincide with the traditional filter when the equations are linear. For a detailed introduction, see [12] or [30].

Given a dynamic process $\mathbf{x}_k$ following a transition equation

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{w}_k) \tag{1}$$

we suppose we have a measurement $\mathbf{z}_k$ such that

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k) \tag{2}$$

where $\mathbf{w}_k$ and $\mathbf{u}_k$ are two mutually-uncorrelated sequences of temporally-uncorrelated Normal random-variables with zero means and covariance matrices $\mathbf{Q}_k$, $\mathbf{R}_k$ respectively[4]. Moreover, $\mathbf{w}_k$ is uncorrelated with $\mathbf{x}_{k-1}$ and $\mathbf{u}_k$ uncorrelated with $\mathbf{x}_k$.
We denote the dimension of $\mathbf{x}_k$ as n_x, the dimension of $\mathbf{w}_k$ as n_w and so on.

We define the *a priori* process estimate as

$$\hat{\mathbf{x}}_k^- = E[\mathbf{x}_k] \tag{3}$$

which is the estimation at time step $k-1$ prior to the step k measurement.
Similarly, we define the *a posteriori* estimate

$$\hat{\mathbf{x}}_k = E[\mathbf{x}_k|\mathbf{z}_k] \tag{4}$$

which is the estimation at time step k after the measurement.

We also have the corresponding estimation errors $\mathbf{e}_k^- = \mathbf{x}_k - \hat{\mathbf{x}}_k^-$ and $\mathbf{e}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k$ and the estimate error covariances

$$\mathbf{P}_k^- = E[\mathbf{e}_k^- \mathbf{e}_k^{-t}]$$
$$\mathbf{P}_k = E[\mathbf{e}_k \mathbf{e}_k^t] \tag{5}$$

where the superscript t corresponds to the transpose operator.

In order to evaluate the above means and covariances we will need the conditional

[4] Some prefer to write $\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{w}_{k-1})$. Needless to say, the two notations are equivalent.

densities $p(\mathbf{x}_k|\mathbf{z}_{k-1})$ and $p(\mathbf{x}_k|\mathbf{z}_k)$, which are determined iteratively via the Time Update and Measurement Update equations.

The *Time Update* step is based upon the Chapman-Kolmogorov equation

$$\begin{aligned} p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) &= \int p(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{z}_{1:k-1})p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})d\mathbf{x}_{k-1} \\ &= \int p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})d\mathbf{x}_{k-1} \end{aligned} \tag{6}$$

via the Markov property.

The *Measurement Update* step is based upon the Bayes rule

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{z}_{1:k-1})}{p(\mathbf{z}_k|\mathbf{z}_{1:k-1})} \tag{7}$$

with

$$p(\mathbf{z}_k|\mathbf{z}_{1:k-1}) = \int p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{z}_{1:k-1})d\mathbf{x}_k$$

A proof of the above is given in the appendix.

EKF is based on the linearization of the transition and measurement equations, and uses the conservation of Normal property within the class of linear functions. We therefore define the Jacobian matrices of $\mathbf{f}$ with respect to the state variable and the system noise as $\mathbf{A}_k$ and $\mathbf{W}_k$ respectively; and similarly for $\mathbf{h}$, as $\mathbf{H}_k$ and $\mathbf{U}_k$ respectively.

More accurately, for every row i and column j we have

$$\mathbf{A}_{ij} = \partial\mathbf{f}_i/\partial\mathbf{x}_j(\hat{\mathbf{x}}_{k-1},0)\ ,\ \mathbf{W}_{ij} = \partial\mathbf{f}_i/\partial\mathbf{w}_j(\hat{\mathbf{x}}_{k-1},0)\ ,\ \mathbf{H}_{ij} = \partial\mathbf{h}_i/\partial\mathbf{x}_j(\hat{\mathbf{x}}_k^-,0)\ ,\ \mathbf{U}_{ij} = \partial\mathbf{h}_i/\partial\mathbf{u}_j(\hat{\mathbf{x}}_k^-,0)$$

Needless to say for a linear system, the function matrices are *equal* to these Jacobians. This is the case for the simple Kalman Filter.

1.1.2 The Algorithm

The actual algorithm could be implemented as follows:

(1) Initialization of $\mathbf{x}_0$ and $\mathbf{P}_0$

For k in $1 \ldots N$

(2) Time Update (Prediction) equations

$$\hat{\mathbf{x}}_k^- = \mathbf{f}(\hat{\mathbf{x}}_{k-1},0) \tag{8}$$

and

$$\mathbf{P}_k^- = \mathbf{A}_k\mathbf{P}_{k-1}\mathbf{A}_k^t + \mathbf{W}_k\mathbf{Q}_{k-1}\mathbf{W}_k^t \tag{9}$$

(3-a) Innovation : We define

$$\hat{\mathbf{z}}_k^- = \mathbf{h}(\hat{\mathbf{x}}_k^-, 0)$$

and

$$\nu_k = \mathbf{z}_k - \hat{\mathbf{z}}_k^-$$

as the innovation process.

(3-b) Measurement Update (Filtering) equations

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k \nu_k \tag{10}$$

and

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)\mathbf{P}_k^- \tag{11}$$

with

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^t (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^t + \mathbf{U}_k \mathbf{R}_k \mathbf{U}_k^t)^{-1} \tag{12}$$

and **I** the Identity matrix.

The above Kalman gain $\mathbf{K}_k$ corresponds to the mean of the conditional distribution of $\mathbf{x}_k$ upon the observation $\mathbf{z}_k$ or equivalently, the matrix that would minimize the mean square error $\mathbf{P}_k$ *within the class of linear estimators*.

This interpretation is based upon the following observation. Having x a Normally distributed random-variable with a mean m_x and variance P_{xx}, and z another Normally distributed random-variable with a mean m_z and variance P_{zz}, and having $P_{zx} = P_{xz}$ the covariance between x and z, the conditional distribution of $x|z$ is also Normal with

$$m_{x|z} = m_x + K(z - m_z)$$

with

$$K = P_{xz} P_{zz}^{-1}$$

which corresponds to our Kalman gain.

1.1.3 Parameter Estimation

For a parameter-set Ψ in the model, the calibration could be carried out via a Maximum Likelihood Estimator (MLE) or in case of conditionally Gaussian state variables, a Quasi-Maximum Likelihood (QML) algorithm.

To do this we need to maximize $\prod_{k=1}^{N} p(z_k|z_{1:k-1})$, and given the Normal form of this probability density function, taking the logarithm, changing the signs and ignoring constant terms, this would be equivalent to minimizing over the parameter-set Ψ

$$L_{1:N} = \sum_{k=1}^{N} ln(P_{z_k z_k}) + \frac{z_k - m_{z_k}}{P_{z_k z_k}} \tag{13}$$

For the KF or EKF we have

$$m_{z_k} = \hat{\mathbf{z}}_k^-$$

and

$$P_{z_k z_k} = \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^t + \mathbf{U}_k \mathbf{R}_k \mathbf{U}_k^t$$

1.2 The Unscented Kalman Filter and Kushner's Nonlinear Filter

1.2.1 Background and Notations

Recently Julier and Uhlmann [18] proposed a new extension of the Kalman Filter to Nonlinear systems, different from the EKF. The new method called the Unscented Kalman Filter (UKF) will calculate the mean to a higher order of accuracy than the EKF, and the covariance to the same order of accuracy.
Unlike the EKF, this method does *not* require any Jacobian calculation since it does not approximate the nonlinear functions of the process and the observation. Indeed it uses the true nonlinear models but approximates the distribution of the state random variable $\mathbf{x}_k$ (as well as the observation $\mathbf{z}_k$) with a Normal distribution by applying an *Unscented Transformation* to it.
In order to be able to apply this Gaussian approximation (unless we have $\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{w}_k$ and $\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{u}_k$, *i.e.* unless the equations are linear in noise) we will need to *augment* the state space by concatenating the noises to it. This augmented state will have a dimension $n_a = n_x + n_w + n_u$.

1.2.2 The Algorithm

The UKF algorithm could be written in the following way:

(1-a) Initialization : Similarly to the EKF, we start with an initial choice for the state vector $\hat{\mathbf{x}}_0 = E[\mathbf{x}_0]$ and its covariance matrix $\mathbf{P}_0 = E[(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^t]$.

We also define the weights $W_i^{(m)}$ and $W_i^{(c)}$ as

$$W_0^{(m)} = \frac{\lambda}{n_a + \lambda}$$

and

$$W_0^{(c)} = \frac{\lambda}{n_a + \lambda} + (1 - \alpha^2 + \beta)$$

and for $i = 1...2n_a$

$$W_i^{(m)} = W_i^{(c)} = \frac{1}{2(n_a + \lambda)} \tag{14}$$

where the scaling parameters α, β, κ and $\lambda = \alpha^2(n_a + \kappa) - n_a$ will be chosen for tuning.

For k in $1...N$

(1-b) State Space Augmentation : As mentioned earlier, we concatenate the state vector with the system noise and the observation noise, and create an augmented state vector for each time-step

$$\mathbf{x}^a_{k-1} = \begin{bmatrix} \mathbf{x}_{k-1} \\ \mathbf{w}_{k-1} \\ \mathbf{u}_{k-1} \end{bmatrix}$$

and therefore

$$\hat{\mathbf{x}}^a_{k-1} = E[\mathbf{x}^a_{k-1}|\mathbf{z}_k] = \begin{bmatrix} \hat{\mathbf{x}}_{k-1} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

and

$$\mathbf{P}^a_{k-1} = \begin{bmatrix} \mathbf{P}_{k-1} & \mathbf{P}_{xw}(k-1|k-1) & \mathbf{0} \\ \mathbf{P}_{xw}(k-1|k-1) & \mathbf{P}_{ww}(k-1|k-1) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{P}_{uu}(k-1|k-1) \end{bmatrix}$$

(1-c) The Unscented Transformation : Following this, in order to use the Normal approximation, we need to construct the corresponding *Sigma Points* through the Unscented Transformation:

$$\chi^a_{k-1}(0) = \hat{\mathbf{x}}^a_{k-1}$$

For $i = 1...n_a$

$$\chi^a_{k-1}(i) = \hat{\mathbf{x}}^a_{k-1} + (\sqrt{(n_a+\lambda)\mathbf{P}^a_{k-1}})_i$$

and for $i = n_a + 1...2n_a$

$$\chi^a_{k-1}(i) = \hat{\mathbf{x}}^a_{k-1} - (\sqrt{(n_a+\lambda)\mathbf{P}^a_{k-1}})_{i-n_a} \tag{15}$$

where the above subscripts i and $i - n_a$ correspond to the i^{th} and $i - n_a^{th}$ columns of the square-root matrix[5].

(2) Time Update equations are

$$\chi_{k|k-1}(i) = \mathbf{f}(\chi^x_{k-1}(i), \chi^w_{k-1}(i)) \tag{16}$$

for $i = 0...2n_a + 1$ and

$$\hat{\mathbf{x}}^-_k = \sum_{i=0}^{2n_a} W_i^{(m)} \chi_{k|k-1}(i) \tag{17}$$

and

$$\mathbf{P}^-_k = \sum_{i=0}^{2n_a} W_i^{(c)} (\chi_{k|k-1}(i) - \hat{\mathbf{x}}^-_k)(\chi_{k|k-1}(i) - \hat{\mathbf{x}}^-_k)^t \tag{18}$$

The superscripts x and w correspond to the state and system-noise portions of the augmented state respectively.

[5] The square-root can be computed via a Cholesky factorization combined with a Singular Value Decomposition. See [27] for details.

(3-a) Innovation : We define

$$\mathbf{Z}_{k|k-1}(i) = \mathbf{h}(\chi_{k|k-1}(i), \chi^u_{k-1}(i)) \tag{19}$$

and

$$\hat{\mathbf{z}}^-_k = \sum_{i=0}^{2n_a} W_i^{(m)} \mathbf{Z}_{k|k-1}(i) \tag{20}$$

and as before

$$\nu_k = \mathbf{z}_k - \hat{\mathbf{z}}^-_k$$

(3-b) Measurement Update

$$\mathbf{P}_{z_k z_k} = \sum_{i=0}^{2n_a} W_i^{(c)} (\mathbf{Z}_{k|k-1}(i) - \hat{\mathbf{z}}^-_k)(\mathbf{Z}_{k|k-1}(i) - \hat{\mathbf{z}}^-_k)^t$$

and

$$\mathbf{P}_{x_k z_k} = \sum_{i=0}^{2n_a} W_i^{(c)} (\chi_{k|k-1}(i) - \hat{\mathbf{x}}^-_k)(\mathbf{Z}_{k|k-1}(i) - \hat{\mathbf{z}}^-_k)^t \tag{21}$$

which gives us the Kalman Gain

$$\mathbf{K}_k = \mathbf{P}_{x_k z_k} \mathbf{P}^{-1}_{z_k z_k}$$

and we have as previously

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}^-_k + \mathbf{K}_k \nu_k \tag{22}$$

as well as

$$\mathbf{P}_k = \mathbf{P}^-_k - \mathbf{K}_k \mathbf{P}_{z_k z_k} \mathbf{K}^t_k \tag{23}$$

which completes the Measurement Update Equations.

1.2.3 Parameter Estimation

The maximization of the likelihood could be done exactly as in (13) taking $\hat{\mathbf{z}}^-_k$ and $P_{z_k z_k}$ defined as above.

1.2.4 Analogy with Kushner's Nonlinear Filter

It would be interesting to compare this algorithm to Kushner's Nonlinear Filter[6] (NLF) based on an approximation of the conditional distribution [19], [20]. In this approach, the authors suggest using a Normal approximation to the densities $p(\mathbf{x}_k|\mathbf{z}_{k-1})$ and $p(\mathbf{x}_k|\mathbf{z}_k)$. They then use the fact that a Normal distribution is entirely determined via its first two moments, which reduces the calculations considerably.
They finally rewrite the moment calculation equations (3), (4) and (5) using the above

[6] This analogy between Kushner's NLF and the UKF has been studied in [16].

$p(\mathbf{x}_k|\mathbf{z}_{k-1})$ and $p(\mathbf{x}_k|\mathbf{z}_k)$, after calculating these conditional densities via the time and measurement update equations (6) and (7). All integrals could be evaluated via Gaussian Quadratures[7].

Note that when $h(x,u)$ is strongly nonlinear, the Gauss Hermite integration is not efficient for evaluating the moments of the measurement update equation, since the term $p(z_k|x_k)$ contains the exponent $z_k - h(x_k)$. The iterative methods based on the idea of importance sampling proposed in [19],[20] correct this problem at the price of a strong increase in computation time. As suggested in [16], one way to avoid this integration would be to make the additional hypothesis that $x_k, h(x_k)|z_{1:k-1}$ is Gaussian.

When $n_x = 1$ and $\lambda = 2$, the numeric integration in the UKF will correspond to a Gauss-Hermite Quadrature of order 3. However in the UKF we can tune the filter and reduce the higher term errors via the previously mentioned parameters α and β.

As the Kushner paper indicates, having an N-dimensional Normal random-variable $\mathbf{X} = \mathcal{N}(\mathbf{m}, \mathbf{P})$ with $\mathbf{m}$ and $\mathbf{P}$ the corresponding mean and covariance, for a polynomial G of degree $2M-1$ we can write

$$E[G(\mathbf{X})] = \frac{1}{(2\pi)^{\frac{N}{2}}|\mathbf{P}|^{\frac{1}{2}}} \int_{R^N} G(\mathbf{y}) exp[-\frac{(\mathbf{y}-\mathbf{m})^t \mathbf{P}^{-1}(\mathbf{y}-\mathbf{m})}{2}] d\mathbf{y}$$

which is equal to

$$E[G(\mathbf{X})] = \sum_{i_1=1}^{M} ... \sum_{i_N=1}^{M} w_{i_1} ... w_{i_N} G(\mathbf{m} + \sqrt{\mathbf{P}}\zeta)$$

where $\zeta^t = \begin{pmatrix} \zeta_{i_1} & ... & \zeta_{i_N} \end{pmatrix}$ is the vector of the Gauss-Hermite roots of order M and $w_{i_1}...w_{i_N}$ are the corresponding weights.
Note that even if both Kushner's NLF and UKF use Gaussian Quadratures, UKF only uses $2N+1$ sigma points, while NLF needs M^N points for the computation of the integrals.

More accurately, for a Quadrature-order M and an N-dimensional (possibly augmented) variable, the sigma-points are defined for $j = 1...N$ and $i_j = 1...M$ as

$$\chi^a_{k-1}(i_1,...,i_N) = \hat{\mathbf{x}}^a_{k-1} + \sqrt{\mathbf{P}^a_{k-1}}\zeta(i_1,...,i_N)$$

where this square-root corresponds to the Cholesky factorization.
Similarly to the UKF, we have the Time Update equations

$$\chi_{k|k-1}(i_1,...,i_N) = \mathbf{f}\left(\chi^x_{k-1}(i_1,...,i_N), \chi^w_{k-1}(i_1,...,i_N)\right)$$

but now

$$\hat{\mathbf{x}}^-_k = \sum_{i_1=1}^{M} ... \sum_{i_N=1}^{M} w_{i_1}...w_{i_N} \chi_{k|k-1}(i_1,...,i_N)$$

[7] A description of the Gaussian Quadrature could be found in [27].

and

$$\mathbf{P}_k^- = \sum_{i_1=1}^{M} \dots \sum_{i_N=1}^{M} w_{i_1} \dots w_{i_N} (\chi_{k|k-1}(i_1,\dots,i_N) - \hat{\mathbf{x}}_k^-)(\chi_{k|k-1}(i_1,\dots,i_N) - \hat{\mathbf{x}}_k^-)^t$$

and similarly for the measurement update equations.

1.3 The Non-Gaussian Case: The Particle Filter

It is possible to generalize the algorithm for the fundamental Gaussian case to one applicable to any distribution. The basic idea is to find the probability density function corresponding to a hidden state x_k at time step k given all the observations $z_{1:k}$ up to that time.

1.3.1 Background and Notations

In this approach, we use Markov-Chain Monte-Carlo simulations instead of using a Gaussian approximation for $(x_k|z_k)$ as the Kalman or Kushner Filters do. A detailed description is given in [10].

The idea is based on the *Importance Sampling* technique:
We can calculate an expected value

$$E[f(x_k)] = \int f(x_k) p(x_k|z_{1:k}) dx_k \tag{24}$$

by using a known and simple *proposal* distribution $q()$.

More precisely, it is possible to write

$$E[f(x_k)] = \int f(x_k) \frac{p(x_k|z_{1:k})}{q(x_k|z_{1:k})} q(x_k|z_{1:k}) dx_k$$

which could be also written as

$$E[f(x_k)] = \int f(x_k) \frac{w_k(x_k)}{p(z_{1:k})} q(x_k|z_{1:k}) dx_k \tag{25}$$

with

$$w_k(x_k) = \frac{p(z_{1:k}|x_k)p(x_k)}{q(x_k|z_{1:k})}$$

defined as the filtering *non-normalized weight* as step k.

We therefore have

$$E[f(x_k)] = \frac{E_q[w_k(x_k)f(x_k)]}{E_q[w_k(x_k)]} = E_q[\tilde{w}_k(x_k)f(x_k)] \tag{26}$$

with

$$\tilde{w}_k(x_k) = \frac{w_k(x_k)}{E_q[w_k(x_k)]}$$

defined as the filtering *normalized weight* as step k.

Using Monte-Carlo sampling from the distribution $q(x_k|z_{1:k})$ we can write in the discrete framework:

$$E[f(x_k)] \approx \sum_{i=1}^{N_{sims}} \tilde{w}_k(x_k^{(i)}) f(x_k^{(i)}) \tag{27}$$

with again

$$\tilde{w}_k(x_k^{(i)}) = \frac{w_k(x_k^{(i)})}{\sum_{j=1}^{N_{sims}} w_k(x_k^{(j)})}$$

Now supposing that our proposal distribution $q()$ satisfies the Markov property, it can be shown that w_k verifies the recursive identity

$$w_k^{(i)} = w_{k-1}^{(i)} \frac{p(z_k|x_k^{(i)})\, p(x_k^{(i)}|x_{k-1}^{(i)})}{q(x_k^{(i)}|x_{k-1}^{(i)}, z_{1:k})} \tag{28}$$

which completes the *Sequential Importance Sampling* algorithm. It is important to note that this means that the state x_k cannot depend on future observations, *i.e.* we are dealing with Filtering and not Smoothing[8].

One major issue with this algorithm is that the variance of the weights increases randomly over time. In order to solve this problem, we could use a *Resampling* algorithm which would map our unequally weighted x_k's to a new set of equally weighted sample points. Different methods have been suggested for this[9].

Needless to say, the choice of the proposal distribution is crucial. Many suggest using

$$q(x_k|x_{k-1}, z_{1:k}) = p(x_k|x_{k-1})$$

since it will give us a simple weight identity

$$w_k^{(i)} = w_{k-1}^{(i)} p(z_k|x_k^{(i)})$$

However this choice of the proposal distribution does *not* take into account our most recent observation z_k at all and therefore could become inefficient.
Hence the idea of using a Gaussian Approximation for the proposal, and in particular an approximation based on the Kalman Filter, in order to incorporate the observations. We therefore will have

$$q(x_k|x_{k-1}, z_{1:k}) = \mathcal{N}(\hat{x}_k, P_k) \tag{29}$$

[8] See [12] for an explanation on Smoothing.
[9] See for instance [4] or [29] for details.

using the same notations as in the section on the Kalman Filter. Such filters are sometimes referred to as the Extended Particle Filter (EPF) and the Unscented Particle Filter (UPF). See [14] for a detailed description of these algorithms.

1.3.2 The Algorithm

Given the above framework, the algorithm for an Extended or Unscented Particle Filter could be implemented in the following way:

(1) For time step $k = 0$ choose x_0 and $P_0 > 0$.
For i such that $1 \leq i \leq N_{sims}$ take

$$x_0^{(i)} = x_0 + \sqrt{P_0} Z^{(i)}$$

where $Z^{(i)}$ is a standard Gaussian simulated number.
Also take $P_0^{(i)} = P_0$ and $w_0^{(i)} = 1/N_{sims}$
While $1 \leq k \leq N$
(2) For each simulation-index i

$$\hat{x}_k^{(i)} = \mathbf{KF}(x_{k-1}^{(i)})$$

with $P_k^{(i)}$ the associated *a posteriori* error covariance matrix.
(KF could be either the EKF or the UKF)
(3) For each i between 1 and N_{sims}

$$\tilde{x}_k^{(i)} = \hat{x}_k^{(i)} + \sqrt{P_k^{(i)}} Z^{(i)}$$

where again $Z^{(i)}$ is a standard Gaussian simulated number.
(4) Calculate the associated weights for each i

$$w_k^{(i)} = w_{k-1}^{(i)} \frac{p(z_k | \tilde{x}_k^{(i)}) p(\tilde{x}_k^{(i)} | x_{k-1}^{(i)})}{q(\tilde{x}_k^{(i)} | x_{k-1}^{(i)}, z_{1:k})}$$

with $q()$ the Normal density with mean $\hat{x}_k^{(i)}$ and variance $P_k^{(i)}$.
(5) Normalize the weights

$$\tilde{w}_k^{(i)} = \frac{w_k^{(i)}}{\sum_{i=1}^{N_{sims}} w_k^{(i)}}$$

(6) Resample the points $\tilde{x}_k^{(i)}$ and get $x_k^{(i)}$ and reset $w_k^{(i)} = \tilde{w}_k^{(i)} = 1/N_{sims}$.
(7) Increment k, Go back to step **(2)** and **Stop** at the end of the **While** loop.

1.3.3 Parameter Estimation

As in the previous section, in order to estimate the parameter-set Ψ we can use an ML Estimator. However since the particle filter does not necessarily assume Gaussian noise, the likelihood function to be maximized has a more general form than the one used in previous sections.
Given the likelihood at step k

$$l_k = p(z_k|z_{1:k-1}) = \int p(z_k|x_k)p(x_k|z_{1:k-1})dx_k$$

the total likelihood is the product of the l_k's above and therefore the log-likelihood to be maximized is

$$ln(L_{1:N}) = \sum_{k=1}^{N} ln(l_k) \tag{30}$$

Now l_k could be written as

$$l_k = \int p(z_k|x_k)\frac{p(x_k|z_{1:k-1})}{q(x_k|x_{k-1},z_{1:k})}q(x_k|x_{k-1},z_{1:k})dx_k$$

and given that by construction the $\tilde{x}_k^{(i)}$'s are distributed according to $q()$, considering the resetting of $w_k^{(i)}$ to a constant $1/N_{sims}$ during the resampling step, we can approximate l_k with

$$\tilde{l}_k = \sum_{i=1}^{N_{sims}} w_k^{(i)}$$

which will provide us with an interpretation of the likelihood as the *total weight*.

3 Stochastic Volatility Models

In this section, we apply the different filters to a few stochastic volatility models including the Heston, the GARCH and the $3/2$ models. To test the performance of each filter, we use five years of S&P500 time-series.
The idea of applying the Kalman Filter to Stochastic Volatility models goes back to Harvey, Ruiz & Shephard [13], where the authors attempt to determine the system parameters via a QML Estimator. This approach has the obvious advantage of simplicity, however it does not account for the nonlinearities and non-Gaussianities of the system. More recently, Pitt & Shephard [10] suggested the use of *Auxiliary* Particle Filters to overcome some of these difficulties. An alternative method based upon the Fourier transform has been presented in [7].

3.1 The State Space Model

Let us first present the state-space form of the stochastic volatility models:

3.1.1 The Heston Model

Let us study the Euler-discretized Heston [15] Stochastic Volatility model in a risk-neutral framework[15]

$$lnS_k = lnS_{k-1} + (r_k - \frac{1}{2}v_k)\Delta t + \sqrt{v_k}\sqrt{\Delta t}B_{k-1} \quad (31)$$

$$v_k = v_{k-1} + (\omega - \theta v_{k-1})\Delta t + \xi\sqrt{v_{k-1}}\sqrt{\Delta t}Z_{k-1} \quad (32)$$

where S_k is the stock price at time-step k, Δt the time interval, r_k the risk-free rate of interest (possibly netted by a dividend-yield), v_k the stock variance and B_k, Z_k two sequences of temporally-uncorrelated Gaussian random-variables with a mutual correlation ρ. The model risk-neutral parameter-set is therefore $\Psi = (\omega, \theta, \xi, \rho)$.

Considering v_k as the hidden state and lnS_{k+1} as the observation, we can subtract from both sides of the transition equation $x_k = f(x_{k-1}, w_{k-1})$, a multiple of the quantity $h(x_{k-1}, u_{k-1}) - z_{k-1}$ which is equal to zero. This would allow us to eliminate the correlation between the system and the measurement noises.
Indeed, if the system equation is

$$v_k = v_{k-1} + (\omega - \theta v_{k-1})\Delta t + \xi\sqrt{v_{k-1}}\sqrt{\Delta t}Z_{k-1} - \rho\xi[lnS_{k-1} + (r_k - \frac{1}{2}v_{k-1})\Delta t + \sqrt{v_{k-1}}\sqrt{\Delta t}B_{k-1} - lnS_k]$$

posing for every k

$$\tilde{Z}_k = \frac{1}{\sqrt{1-\rho^2}}(Z_k - \rho B_k)$$

[15]The same exact methodology could be used in a non risk-neutral setting. We are supposing we have a small time-step Δt in order to be able to apply the Girsanov theorem.

we will have as expected $\tilde{Z}_k$ *uncorrelated* with B_k and

$$x_k = v_k = v_{k-1} + [\omega - \rho\xi r_k - (\theta - \frac{1}{2}\rho\xi)v_{k-1}]\Delta t + \rho\xi ln(\frac{S_k}{S_{k-1}}) + \xi\sqrt{1-\rho^2}\sqrt{v_{k-1}}\sqrt{\Delta t}\tilde{Z}_{k-1} \tag{33}$$

and the measurement equation would be

$$z_k = lnS_{k+1} = lnS_k + (r_k - \frac{1}{2}v_k)\Delta t + \sqrt{v_k}\sqrt{\Delta t}B_k \tag{34}$$

3.1.2 Other Stochastic Volatility Models

It is easy to generalize the above state space model to other stochastic volatility approaches. Indeed we could replace (32) with

$$v_k = v_{k-1} + (\omega - \theta v_{k-1})\Delta t + \xi v_{k-1}^p\sqrt{\Delta t}Z_{k-1} \tag{35}$$

where $p = 1/2$ would naturally correspond to the Heston (Square-Root) model, $p = 1$ to the GARCH diffusion-limit model, and $p = 3/2$ to the $3/2$ model. These models have all been described and analyzed in [23].
The new state transition equation would therefore become

$$v_k = v_{k-1} + \left[\omega - \rho\xi r_k v_{k-1}^{p-\frac{1}{2}} - \left(\theta - \frac{1}{2}\rho\xi v_{k-1}^{p-\frac{1}{2}}\right)v_{k-1}\right]\Delta t + \rho\xi v_{k-1}^{p-\frac{1}{2}} ln(\frac{S_k}{S_{k-1}}) + \xi\sqrt{1-\rho^2}v_{k-1}^p\sqrt{\Delta t}\tilde{Z}_{k-1} \tag{36}$$

where the same choice of state space $x_k = v_k$ is made.

3.1.3 Robustness and Stability

In this state-space formulation, we only need to choose a value for v_0 which could be set to the historic-volatility over a period preceding our time-series. Ideally, the choice of v_0 should not affect the results enormously, *i.e.* we should have a robust system.

As we saw in the previous section, the system stability greatly depends on the measurement noise. However in this case the system noise is precisely $\sqrt{\Delta t}v_k$, and therefore we do not have a direct control on this issue. Nevertheless we could add an independent Normal measurement noise with a given variance R

$$z_k = lnS_{k+1} = lnS_k + (r_k - \frac{1}{2}v_k)\Delta t + \sqrt{v_k}\sqrt{\Delta t}B_k + RW_k$$

which would allow us to *tune* the filter[16].

[16] An alternative approach would consist in choosing a volatility *proxy* such as $f(lnS_k, lnS_{k+1}) = lnS_{k+1} - lnS_k$ and ignoring the stock-drift term, given that $\Delta t = o(\sqrt{\Delta t})$. We would therefore write $z_k = ln(|f(lnS_k, lnS_{k+1})|) = E[ln(|f(lnS_k^*, lnS_{k+1}^*)|)] + \frac{1}{2}lnv_k + \varepsilon_k$ with S_t^* the same process as S_t but with a volatility of *one*, and ε_k corresponding to the measurement noise. See [2] for details.

3.2 The Filters

We can now apply the Gaussian and the Particle Filters to our problem:

3.2.1 Gaussian Filters

For the EKF we will have

$$A_k = 1 - \left[\rho\xi r_k(p-\frac{1}{2})v_{k-1}^{p-\frac{3}{2}} + \theta - \frac{1}{2}\rho\xi(p+\frac{1}{2})v_{k-1}^{p-\frac{1}{2}}\right]\Delta t + (p-\frac{1}{2})\rho\xi v_{k-1}^{p-\frac{3}{2}} ln(\frac{S_k}{S_{k-1}})$$

and

$$W_k = \xi\sqrt{1-\rho^2}v_{k-1}^{p}\sqrt{\Delta t}$$

as well as

$$H_k = -\frac{1}{2}\Delta t$$

and

$$U_k = \sqrt{v_k}\sqrt{\Delta t}$$

The same time update and measurement update equations could be used with the UKF or Kushner's NLF.

3.2.2 Particle Filters

We could also apply the Particle Filtering algorithm to our problem. Using the same notations as in section 1.3.2 and calling

$$\mathbf{n}(x,\mathbf{m},\mathbf{s}) = \frac{1}{\sqrt{2\pi}\mathbf{s}} exp(-\frac{(x-\mathbf{m})^2}{2\mathbf{s}^2})$$

the Normal density with mean **m** and standard deviation **s**, we will have

$$q(\tilde{x}_k^{(i)}|x_{k-1}^{(i)}, z_{1:k}) = \mathbf{n}\left(\tilde{x}_k^{(i)}, \mathbf{m} = \hat{x}_k^{(i)}, \mathbf{s} = \sqrt{P_k^{(i)}}\right)$$

as well as

$$p(z_k|\tilde{x}_k^{(i)}) = \mathbf{n}\left(z_k, \mathbf{m} = z_{k-1} + (r_k - \frac{1}{2}\tilde{x}_k^{(i)})\Delta t, \mathbf{s} = \sqrt{\tilde{x}_k^{(i)}}\sqrt{\Delta t}\right)$$

and

$$p(\tilde{x}_k^{(i)}|x_{k-1}^{(i)}) = \mathbf{n}\left(\tilde{x}_k^{(i)}, \mathbf{m_x}, \mathbf{s} = \xi\sqrt{1-\rho^2}(x_{k-1}^{(i)})^p\sqrt{\Delta t}\right)$$

with

$$\mathbf{m_x} = x_{k-1}^{(i)} + \left[\omega - \rho\xi r_k(x_{k-1}^{(i)})^{p-\frac{1}{2}} - \left(\theta - \frac{1}{2}\rho\xi(x_{k-1}^{(i)})^{p-\frac{1}{2}}\right)x_{k-1}^{(i)}\right]\Delta t + \rho\xi(x_{k-1}^{(i)})^{p-\frac{1}{2}}(z_{k-1} - z_{k-2})$$

and as before we have

$$w_k^{(i)} = w_{k-1}^{(i)} \frac{p(z_k|\tilde{x}_k^{(i)})p(\tilde{x}_k^{(i)}|x_{k-1}^{(i)})}{q(\tilde{x}_k^{(i)}|x_{k-1}^{(i)}, z_{1:k})}$$

which provides us with what we need for the filter implementation.

3.3 Parameter Estimation and Back-Testing

For the Gaussian MLE we will need to minimize

$$\phi(\omega,\theta,\xi,\rho) = \sum_{k=1}^{K}[ln(F_k) + \frac{\nu_k^2}{F_k}]$$

with $\nu_k = \mathbf{z}_k - \hat{\mathbf{z}}_k^-$ and $\mathbf{F}_k = \mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^t + \mathbf{U}_k\mathbf{U}_k^t$.

For the Particle MLE, as previously mentioned, we need to maximize

$$\sum_{k=1}^{N} ln\left(\sum_{i=1}^{N_{sims}} w_k^{(i)}\right)$$

By maximizing the above likelihood functions, we will find the optimal parameter-set $\hat{\Psi} = (\hat{\omega},\hat{\theta},\hat{\xi},\hat{\rho})$. This calibration procedure could then be used for pricing of derivatives instruments or forecasting volatility.

We could perform a back-testing process in the following way. Choosing an original parameter-set

$$\Psi^* = (0.02, 0.5, 0.05, -0.5)$$

and using a Monte-Carlo simulation, we generate an artificial time-series. We take $S_0 = 1000$ USD, $v_0 = 0.04$, $r_k = 0.027$ and $\Delta t = 1$. Note that we are taking a large Δt in order to have meaningful errors. The time-series is generated via the above transition equation (32) and the usual log-normal measurement equation.
We find the following optimal[17] parameter-sets:

$$\hat{\Psi}_{EKF} = (0.036335, 0.928632, 0.036008, -1.000000)$$
$$\hat{\Psi}_{UKF} = (0.033104, 0.848984, 0.033263, -0.983985)$$
$$\hat{\Psi}_{EPF} = (0.019357, 0.500021, 0.033354, -0.581794)$$
$$\hat{\Psi}_{UPF} = (0.019480, 0.489375, 0.047030, -0.229242)$$

which shows the better performance of the Particle Filters. However, it should be reminded that the Particle Filters are also more computation-intensive than the Gaussian ones.

[17] In this section, all optimizations were made via the Direction-Set algorithm as described in [27]. The precision was set to 1.0e-6.

3.4 Application to the S&P500 Index

The above filters were applied to five years of S&P500 time-series (1996 to 2001) in [17] and the filtering errors were considered for the Heston model, the GARCH model and the $3/2$ model. Daily index close-prices were used for this purpouse, and the time interval was set to $\Delta t = 1/252$. The appropriate risk-free rate was applied and was adjusted by the index dividend yield at the time of the measurement.

As in the previous section, the performance could be measured via the MPE and the RMSE. We could also refer to figures 5 to 11 for a visual interpretation of the performance measurements.

$$
\begin{aligned}
MPE_{EKF-Heston} &= 3.58207e-05 & RMSE_{EKF-Heston} &= 1.83223e-05 \\
MPE_{EKF-GARCH} &= 2.78438e-05 & RMSE_{EKF-GARCH} &= 1.42428e-05 \\
MPE_{EKF-\frac{3}{2}} &= 2.63227e-05 & RMSE_{EKF-\frac{3}{2}} &= 1.74760e-05
\end{aligned}
$$

$$
\begin{aligned}
MPE_{UKF-Heston} &= 3.00000e-05 & RMSE_{UKF-Heston} &= 1.91280e-05 \\
MPE_{UKF-GARCH} &= 2.99275e-05 & RMSE_{UKF-GARCH} &= 2.58131e-05 \\
MPE_{UKF-\frac{3}{2}} &= 2.82279e-05 & RMSE_{UKF-\frac{3}{2}} &= 1.55777e-05
\end{aligned}
$$

$$
\begin{aligned}
MPE_{EPF-Heston} &= 2.70104e-05 & RMSE_{EPF-Heston} &= 1.34534e-05 \\
MPE_{EPF-GARCH} &= 2.48733e-05 & RMSE_{EPF-GARCH} &= 4.99337e-06 \\
MPE_{EPF-\frac{3}{2}} &= 2.26462e-05 & RMSE_{EPF-\frac{3}{2}} &= 2.58645e-06
\end{aligned}
$$

$$
\begin{aligned}
MPE_{UPF-Heston} &= 2.04000e-05 & RMSE_{UPF-Heston} &= 2.74818e-06 \\
MPE_{UPF-GARCH} &= 2.63036e-05 & RMSE_{UPF-GARCH} &= 8.44030e-07 \\
MPE_{UPF-\frac{3}{2}} &= 1.73857e-05 & RMSE_{UPF-\frac{3}{2}} &= 4.09918e-06
\end{aligned}
$$

Two immediate observations can be made: On the one hand the Particle Filters have a better performance than the Gaussian ones, which reconfirms what one would anticipate. On the other hand for most of the Filters, the $3/2$ model seems to outperform the Heston model, which is in line with the findings of Engle & Ishida [11].

3.5 Conclusion

Using the Gaussian or Particle Filtering techniques, it is possible to estimate the stochastic volatility parameters from the underlying asset time-series, in the risk-neutral or real-world context.
As expected, we observe an improvement when Particle Filters are used. What is more,

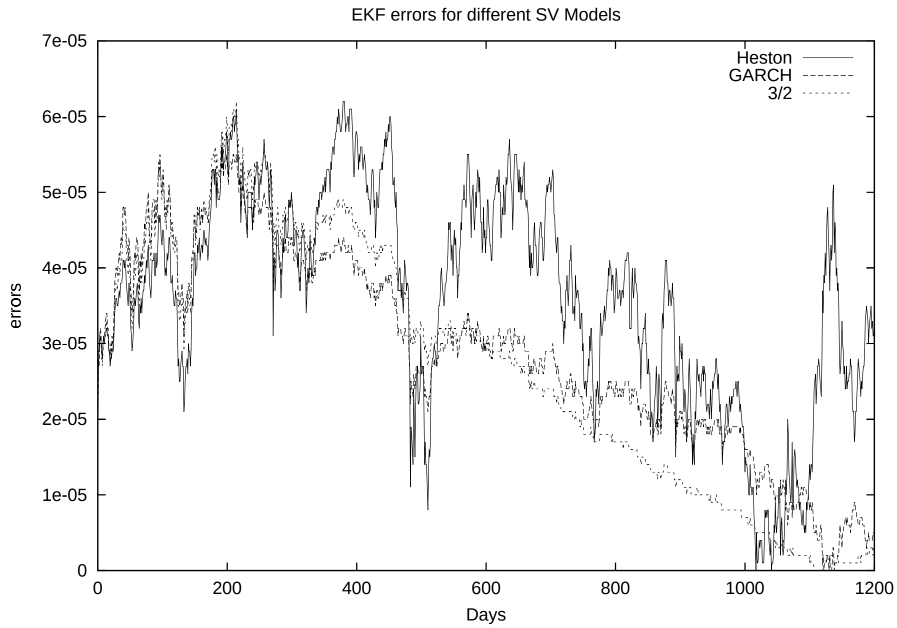

Figure 5: Comparison of EKF Filtering errors for Heston, GARCH and 3/2 Models.

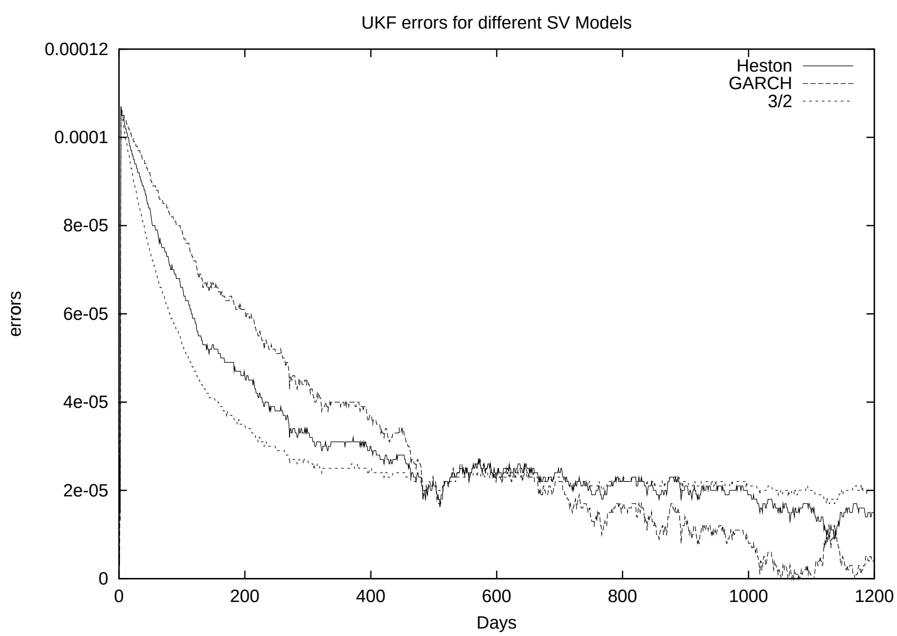

Figure 6: Comparison of UKF Filtering errors for Heston, GARCH and 3/2 Models.

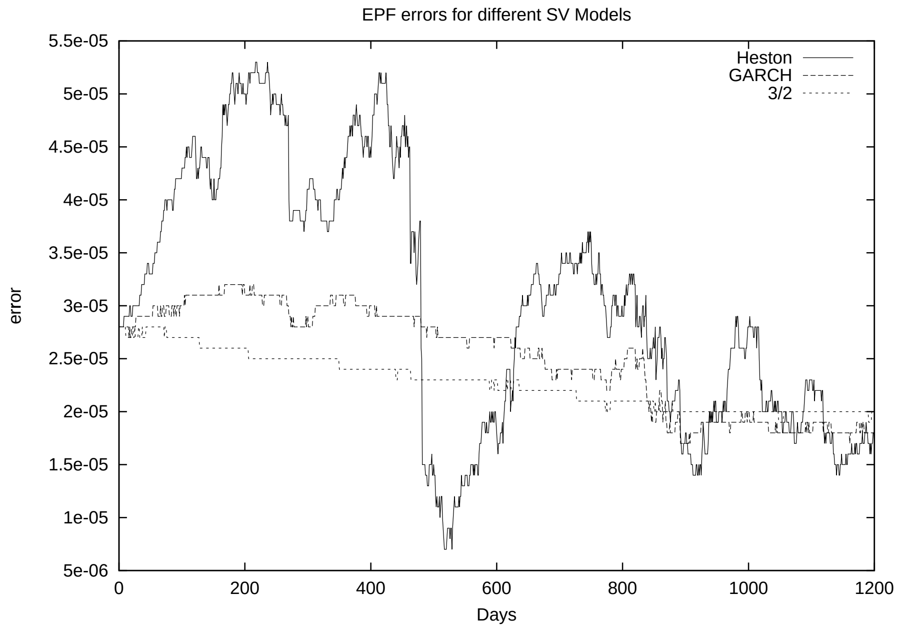

Figure 7: Comparison of EPF Filtering errors for Heston, GARCH and 3/2 Models.

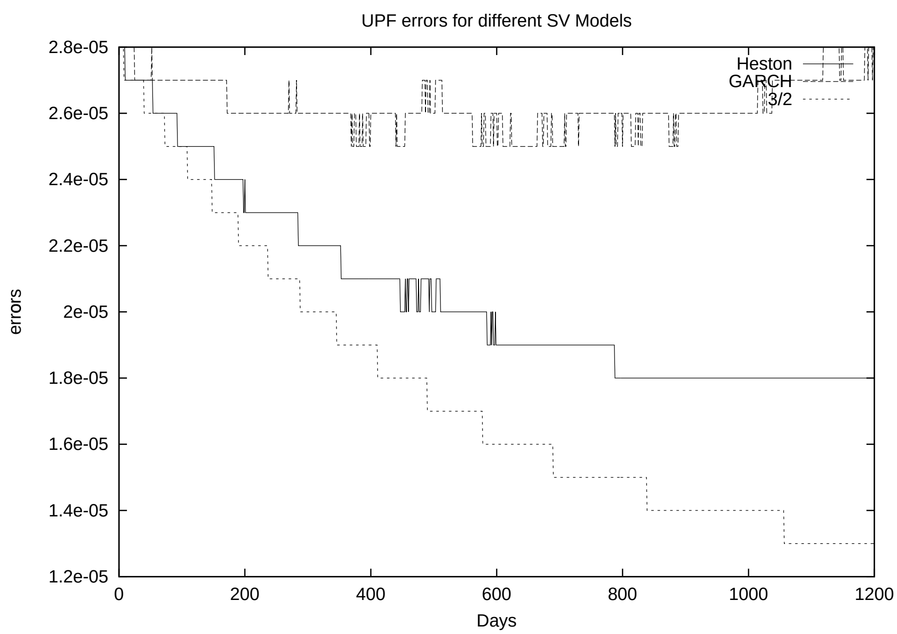

Figure 8: Comparison of UPF Filtering errors for Heston, GARCH and 3/2 Models.

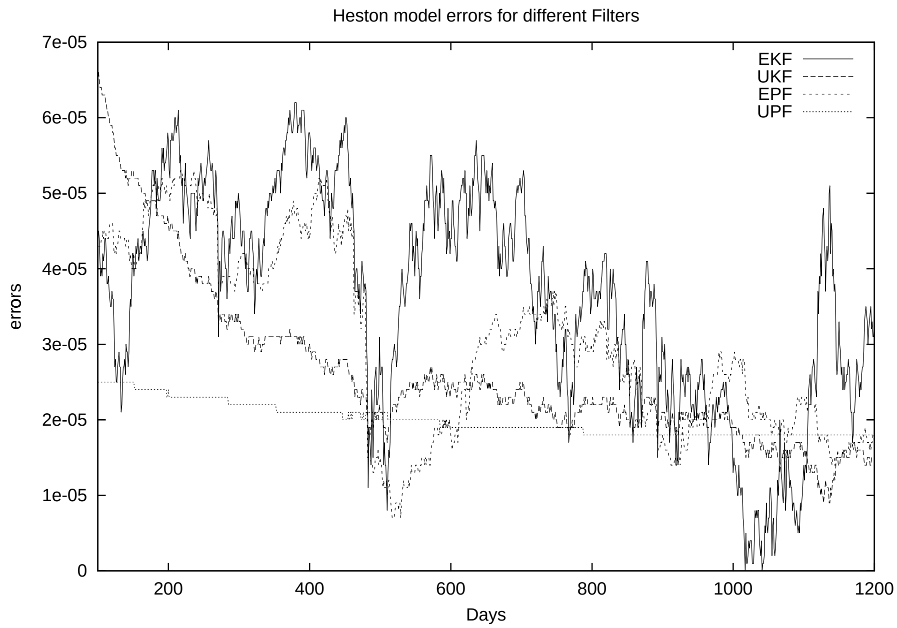

Figure 9: Comparison of Filtering errors for the Heston Model.

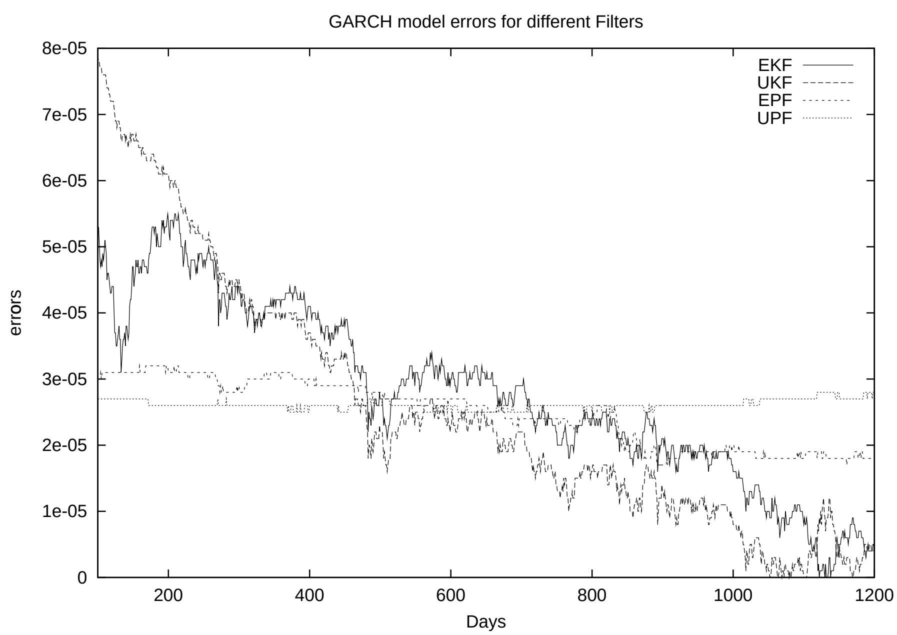

Figure 10: Comparison of Filtering errors for the GARCH Model.

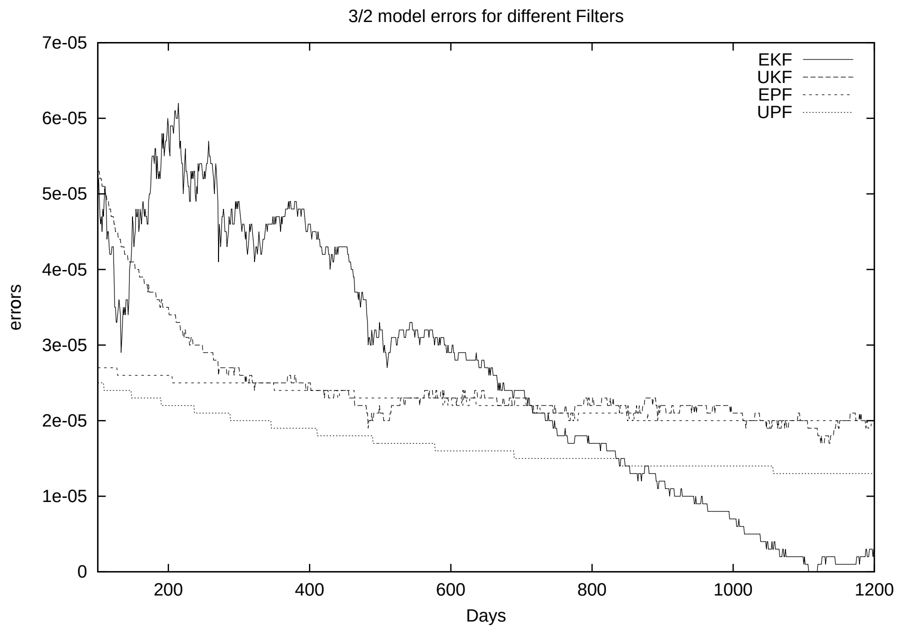

Figure 11: Comparison of Filtering errors for the 3/2 Model.

given the tests carried out on the S&P500 data it seems that, despite its vast popularity, the Heston model does not perform as well as the $3/2$ representation.
This suggests further research on other existing models such as Jump Diffusion [25], Variance Gamma [24] or CGMY [9]. Clearly, because of the non-Gaussianity of these models, the Particle Filtering technique will need to be applied to them[18].
Finally it would be instructive to compare the risk-neutral parameter-set obtained from the above *time-series* based approaches, to the parameter-set resulting from a *cross-sectional* approach using options market prices at a given point in time[19]. Inconsistent parameters between the two approaches would signal either arbitrage opportunities in the market, or a misspecification in the tested model[20].

[18] A study on Filtering and Lévy processes has recently been done in [6].
[19] This idea is explored in [1] in a Non-parametric fashion.
[20] This comparison supposes that the Girsanov theorem *is* applicable to the tested model.

4 Summary

In this article, we present an introduction to various filtering algorithms: the Kalman filter, the Extended Filter, as well as the Unscented Kalman Filter (UKF) similar to Kushner's Nonlinear filter. We also tackle the subject of Non-Gaussian filters and describe the Particle Filtering (PF) algorithm.

We then apply the filters to a term structure model of commodity prices. Our main results are the following: Firstly, the approximation introduced in the Extended filter has an influence on the model performances. Secondly, the estimation results are sensitive to the system matrix containing the errors of the measurement equation. Thirdly, the approximation made in the extended filter is not a real issue until the model becomes highly nonlinear. In that case, other nonlinear filters such as those described in section 1.2 may be used.

Lastly, the application of the filters to stochastic volatility models shows that the Particle Filters perform better than the Gaussian ones, however they are also more expensive. What is more, given the tests carried out on the S&P500 data it seems that, despite its vast popularity, the Heston model does not perform as well as the $3/2$ representation. This suggests further research on other existing models such as Jump Diffusion, Variance Gamma or CGMY. Clearly, because of the non-Gaussianity of these models, the Particle Filtering technique will need to be applied to them.

Appendix

The Measurement Update equation is

$$p(x_k|z_{1:k}) = \frac{p(z_k|x_k)\,p(x_k|z_{1:k-1})}{p(z_k|z_{1:k-1})}$$

where the denominator $p(z_k|z_{1:k-1})$ could be written as

$$p(z_k|z_{1:k-1}) = \int p(z_k|x_k)\,p(x_k|z_{1:k-1})dx_k$$

and corresponds to the Likelihood Function for the time-step k.

Indeed using the Bayes rule and the Markov property, we have

$$\begin{aligned}
p(x_k|z_{1:k}) &= \frac{p(z_{1:k}|x_k)\,p(x_k)}{p(z_{1:k})} \\
&= \frac{p(z_k, z_{1:k-1}|x_k)\,p(x_k)}{p(z_k, z_{1:k-1})} \\
&= \frac{p(z_k|z_{1:k-1}, x_k)\,p(z_{1:k-1}|x_k)\,p(x_k)}{p(z_k|z_{1:k-1})\,p(z_{1:k-1})} \\
&= \frac{p(z_k|z_{1:k-1}, x_k)\,p(x_k|z_{1:k-1})\,p(z_{1:k-1})\,p(x_k)}{p(z_k|z_{1:k-1})\,p(z_{1:k-1})\,p(x_k)} \\
&= \frac{p(z_k|x_k)\,p(x_k|z_{1:k-1})}{p(z_k|z_{1:k-1})} \qquad \square
\end{aligned}$$

Note that $p(x_k|z_k)$ is proportinal to

$$exp\left(-\frac{1}{2}(z_k - h(x_k))^t R_k^{-1}(z_k - h(x_k))\right)$$

under the hypothesis of additive measurement noises.

References

[1] Aït-Sahalia Y., Wang Y., Yared F. (2001) "Do Option Markets Correctly Price the Probabilities of Movement of the Underlying Asset?" *Journal of Econometrics*, 101

[2] Alizadeh S., Brandt M.W., Diebold F.X. (2002) "Range-Based Estimation of Stochastic Volatility Models" *Journal of Finance*, Vol. 57, No. 3

[3] Anderson B. D. O., Moore J. B. (1979) "Optimal Filtering" *Englewood Cliffs, Prentice Hall*

[4] Arulampalam S., Maskell S., Gordon N., Clapp T. (2002) "A Tutorial on Particle Filters for On-line Nonlinear/ Non-Gaussian Bayesian Tracking" *IEEE Transactions on Signal Processing*, Vol. 50, No. 2

[5] Babbs S. H., Nowman K. B. (1999) "Kalman Filtering of Generalized Vasicek Term Structure Models" *Journal of Financial and Quantitative Analysis*, Vol. 34, No. 1

[6] Barndorff-Nielsen O. E., Shephard N. (2002) "Econometric Analysis of Realized Volatility and its use in Estimating Stochastic Volatility Models" *Journal of the Royal Statistical Society*, Series B, Vol. 64

[7] Bates D. S. (2002) "Maximum Likelihood Estimation of Latent Affine Processes" *University of Iowa & NBER*

[8] Brennan M.J., Schwartz E.S. (1985) "Evaluating Natural Resource Investments" *The Journal of Business*, Vol. 58, No. 2

[9] Carr P., Geman H., Madan D., Yor M. (2002) "The Fine Structure of Asset Returns" *Journal of Business*, Vol. 75, No. 2

[10] Doucet A., De Freitas N., Gordon N. *(Editors)* (2001) "Sequential Monte-Carlo Methods in Practice" *Springer-Verlag*

[11] Engle R. F., Ishida I. (2002) "The Square-Root, the Affine, and the CEV GARCH Models" *Working Paper, New York University and University of California San Diego*

[12] Harvey A. C. (1989) "Forecasting, Structural Time Series Models and the Kalman Filter" *Cambridge University Press*

[13] Harvey A. C., Ruiz E., Shephard N. (1994) "Multivariate Stochastic Variance Models" *Review of Economic Studies*, Vol. 61, No. 2

[14] Haykin S. *(Editor)* (2001) "Kalman Filtering and Neural Networks" *Wiley Inter-Science*

[15] Heston S. (1993) "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options" *Review of Financial Studies*, Vol. 6, No. 2

[16] Ito K., Xiong K. (2000) "Gaussian Filters for Nonlinear Filtering Problems" *IEEE Transactions on Automatic Control*, Vol. 45, No. 5

[17] Javaheri A. (2002) "The Volatility Process" *Ph.D. Thesis in progress, Ecole des Mines de Paris*

[18] Julier S. J., Uhlmann J.K. (1997) "A New Extension of the Kalman Filter to Nonlinear Systems" *The University of Oxford, The Robotics Research Group*

[19] Kushner H. J. (1967) "Approximations to Optimal Nonlinear Filters" *IEEE Transactions on Automatic Control*, Vol. 12

[20] Kushner H. J., Budhiraja A. S. (2000) "A Nonlinear Filtering Algorithm based on an Approximation of the Conditional Distribution" *IEEE Transactions on Automatic Control*, Vol. 45, No. 3

[21] Lautier D. (2000) "La Structure par Terme des Prix des Commodités : Analyse Théorique et Applications au Marché Pétrolier" *Thèse de Doctorat, Université Paris IX*

[22] Lautier D., Galli A. (2000) "Un Modèle de Structure par Terme des Prix des Matières Premières avec Comportement Asymétrique du Rendement d'Opportunité" *Finéco*, Vol. 10

[23] Lewis A. (2000) "Option Valuation under Stochastic Volatility" *Finance Press*

[24] Madan D., Carr P., Chang E.C. (1998) "The Variance Gamma Process and Option Pricing" *European Finance Review*, Vol. 2, No. 1

[25] Merton R. C. (1976) "Option Pricing when the Underlying Stock Returns are Discontinuous" *Journal of Financial Economics*, No. 3

[26] Pennacchi G. G. (1991) "Identifying the Dynamics of Real Interest Rates and Inflation : Evidence Using Survey Data" *The Review of Financial Studies*, Vol. 4, No. 1

[27] Press W. H., Teukolsky S.A., Vetterling W.T., Flannery B. P. (1997) "Numerical Recipes in C: The Art of Scientific Computing" *Cambridge University Press*, 2nd Edition

[28] Schwartz E.S. (1997) "The Stochastic Behavior of Commodity Prices : Implications for Valuation and Hedging" *The Journal of Finance*, Vol. 52, No. 3

[29] Van der Merwe R., Doucet A., de Freitas N., Wan E. (2000) "The Unscented Particle Filter" *Oregon Graduate Institute, Cambridge University and UC Berkeley*

[30] Welch G., Bishop G. (2002) "An Introduction to the Kalman Filter" *Department of Computer Science, University of North Carolina at Chapel Hill*

[31] Wells C. (1996) "The Kalman Filter in Finance" *Advanced Studies in Theoretical and Applied Econometrics, Kluwer Academic Publishers*, Vol. 32