Sequential calibration of options

Erik Lindström, Jonas Ströjby, Mats Brodén, Magnus Wiktorsson*, Jan Holst

Division of Mathematical Statistic, Centre for Mathematical Sciences, Lund University, Box 118, SE-22100 Lund, Sweden

Available online 29 August 2007

Abstract

Robust calibration of option valuation models to quoted option prices is non-trivial but crucial for good performance. A framework based on the state-space formulation of the option valuation model is introduced. Non-linear (Kalman) filters are needed to do inference since the models have latent variables (e.g. volatility). The statistical framework is made adaptive by introducing stochastic dynamics for the parameters. This allows the parameters to change over time, while treating the measurement noise in a statistically consistent way and using all data efficiently. The performance and computational efficiency of standard and iterated extended Kalman filters (EKF and IEKF) are investigated. These methods are compared to common calibration such as weighted least squares (WLS) and penalized weighted least squares (PWLS). A simulation study, using the Bates model, shows that the adaptive framework is capable of tracking time varying parameters and latent processes such as stochastic volatility processes. It is found that the filter estimates are the most accurate, followed by the PWLS estimates. The estimates from all of the advanced methods are significantly closer to the true parameters than the WLS estimates which overfits data. The filters are also faster than least squares methods. All calibration methods are also applied to daily European option data on the S&P 500 index, where the Heston, Bates and NIG-CIR models are considered. The results are similar to the simulation study and it can be seen that the overfitting is a real problem for the WLS estimator when applied complex models.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Non-linear Kalman filters; Calibration; Option valuation

1. Introduction

The predictive power of parametric option valuation models was recognized immediately after the [Black and Scholes \(1973\)](#) model was introduced, cf. [Campbell et al. \(1997\)](#). More recent models have included jumps, see [Merton \(1976\)](#), stochastic volatility, see e.g. [Hull and White \(1987\)](#) and [Heston \(1993\)](#) and state dependent diffusion terms, see e.g. [Dupire \(1994\)](#) and [Derman and Kani \(1994\)](#) in order to improve the predictive power further. Today, models use a combination of jumps, stochastic volatility and local volatility, see e.g. [Bates \(1996\)](#), [Duffie et al. \(2000\)](#), [Carr et al. \(2003b\)](#), [Chernov et al. \(2004\)](#), and [Carr and Wu \(2004\)](#). The purpose of this paper is to outline a framework in which these advanced models can be adaptively calibrated to data.

There are some issues when working with option data. A serious problem is the absence of a single quoted market price. Instead, ask and bid prices are quoted and it is common practice to approximate the market price by the

* Corresponding author. Fax: +46 46 2224623.

E-mail address: magnusw@maths.lth.se (M. Wiktorsson).

mid-price of these. Using the mid-price increases the risk for overfitting as any parameter vector yielding option values within the bid–ask spread is acceptable with respect to the arbitrage conditions. The overfitting is accentuated when using state of the art models as complex models are more prone to overfitting than simpler, transparent models.

Another difficulty arises from the fact that market properties change over time. This can be e.g. macro-economic events as well as investor preferences, corresponding to time varying $\mathbb{P}$ and $\mathbb{Q}$ measures, respectively. Time varying parameters make adaptive calibration necessary. Less complex models are easier to adapt to changing conditions and are therefore often used.

Models can be calibrated using data mining methods if the main purpose is to interpolate or predict prices of contracts of the same type as the calibration instruments. However, good parameter estimates are needed if the purpose of the calibration is applications like hedging, risk management or pricing exotics. The most common calibration technique is some version of daily least squares estimation, which provides good in sample predictions but is also known to be notoriously non-robust to outliers giving bad parameter estimates, cf. Cont and Tankov (2004). A problem frequently associated with the least squares calibration methods is multimodal loss functions, making numerical optimization difficult. Several algorithms designed to reduce this problem by adding penalties have been suggested. However, the main purpose of these is to improve the numerical stability, not the quality of the parameter estimates.

Adaptive estimators are rare in the academic literature but commonly used in the industry, often using some ad hoc adjustment. The dominating approach is to estimate the parameters using rolling non-linear regression techniques, often in combination with some penalty.

This paper suggests that standard least squares methods should be replaced by a non-linear (Kalman) filter method, as this decomposes the residual error into changes in the latent variables (e.g. volatility) and pricing errors. The Kalman filter will automatically use all data but assign more weight to recent data in a statistically sound way.

The method is made adaptive by introducing the parameters as latent states, and assigning dynamics to these. Filtering the augmented latent process will estimate the original latent state and parameters simultaneously. We find that the proposed filter method is both faster than and at least as accurate as the least squares methods on the estimation and validation set and significantly more accurate when pricing other contracts.

2. Option valuation

It is essential that the theoretical value of the calibration instruments (i.e. the options used in the calibration) can be calculated with high accuracy using limited computational resources. These requirements rule out many quoted options, essentially restricting the available instruments to European call and put options. The values of these instruments can be obtained using Fourier methods, cf. Carr and Madan (1999).

The Fourier transform based techniques are both fast and fairly accurate. The main limitation is that we need to know the characteristic function for the log stock price in closed form. However, for a fairly large class of stock price models e.g. Black–Scholes (Black and Scholes, 1973), Merton (Merton, 1976), Heston (Heston, 1993), Bates (Bates, 1996), Normal inverse Gaussian (NIG) (see e.g. Barndorff-Nielsen, 1997 and the references therein), variance Gamma (VG) (Madan and Seneta, 1990), BNS (Barndorff-Nielsen and Shepard, 2001, 2002), CGMY (Carr et al., 2002), finite moment log-stable (FMLS) (Carr and Wu, 2003), NIG-CIR model (Carr et al., 2003b), a class of time-changed exponential Lévy models (Carr and Wu, 2004; Huang and Wu, 2004), etc. the characteristic function is available in closed form. We will consider some of these models in detail below (cf. Section 2.1). The Fourier transform based methods have been used in financial mathematics for some time now (see Carr and Madan, 1999) for a good reference to start with. A long list of references to articles using Fourier transform based methods can be found in Carr et al. (2003a).

2.1. Stock price models

In this paper we consider the Bates model (Bates, 1996) and a sub-model (cf. below) as well as the NIG-CIR model (cf. below). This choice of models is based on the results in Schoutens et al. (2004).

2.1.1. The Bates model and its sub-models

The dynamics for the Bates model under the risk-neutral measure $\mathbb{Q}$ is given by

$$dS_t = rS_t dt + \sqrt{V_t} S_t dW_t^{(S)} + S_t dJ_t, \quad (1)$$

$$dV_t = \kappa(\xi - V_t) dt + \sigma_V \sqrt{V_t} dW_t^{(V)}, \quad (2)$$

where $W^{(S)}$ and $W^{(V)}$ are standard Brownian motions with correlation ρ , κ is the mean reversion rate, ξ is the mean reversion level and with V_0 as the initial value of V . Further, J is a compound Poisson process with intensity λ , having jumps ΔJ such that $\log(1 + \Delta J) \in N(\mu, \delta^2)$ and drift $-\lambda(\exp(\delta^2/2 + \mu) - 1)$. This choice of drift for J forces the discounted stock price process to be a martingale under $\mathbb{Q}$. The Bates model contains several models (Black–Scholes, Merton and Heston) as sub-models.

2.1.2. The NIG-CIR model

The NIG-CIR model is a Normal inverse Gaussian model that is stochastically time shifted by an integrated Cox–Ingersoll–Ross stochastic volatility process. This model was first described in Carr et al. (2003b). Here the stock price process

$$S_t = S_0 \exp(X_{I_t}), \quad (3)$$

where X_t is a NIG Lévy process, having parameters δ, α, β and where $I_t = \int_0^t V_s ds$, V_t has dynamics according to Eq. (2). Here the process $\{V_t\}_{t \geq 0}$ is assumed to be independent of the process $\{X_t\}_{t \geq 0}$. It is, however, possible to calculate the moment generating function without this assumption (see Huang and Wu, 2004; Carr and Wu, 2004).

2.2. Implementation and numerical aspects of the inverse Fourier transform

Lee (2004) treats error bounds for a fast Fourier transform (FFT) implementation of the Fourier transform method and list the generalized Fourier transforms for some common pay-off functions such as e.g. the European call option. Using FFT, one can obtain prices on a regularly spaced grid of log strike levels. Options observed on the market are usually regularly spaced in the strike level, not the log strike level. If we are to evaluate the prices for a set of European call options we need to fit the corresponding log strike levels into a regular spaced grid. The accuracy of FFT depends on the grid spacing used, the interval on which integrand is evaluated and the properties of the characteristic function of the log stock price. One FFT is needed for each time to maturity. This will lead to extra work if we have several dates of maturity each with only a few strike levels. In order to avoid using different grids for different strike levels and maturities we will here instead consider a Gauss–Laguerre quadrature implementation for the inverse Fourier transform.

In order to price a European call option we need to calculate the inverse Fourier transform (Lee, 2004):

$$C_t(K, T - t) = p(t, T) \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-ik\omega - \gamma k} e^{s_t(1 + \gamma + i\omega)} M_{X_{T|t}}(1 + \gamma + i\omega)}{(\gamma + i\omega)(1 + \gamma + i\omega)} d\omega, \quad (4)$$

where t is the present time, T time of maturity, $K = \exp(k)$ is the strike level, $S_t = \exp(s_t)$ is the current dividend corrected stock level (cf. Section 5 for precise details), $M_{X_{T|t}}$ is the moment generating function of $X_T = \log(S_T/S_t)$ under the risk neutral measure $\mathbb{Q}$ given S_t and $p(t, T)$ is a discounting factor for the time interval $[t, T]$. This is well defined if there exists a positive real number γ such that

$$E^{\mathbb{Q}}[S_T^{1+\gamma}] < \infty.$$

In order to price a Binary call option we instead need to calculate the inverse Fourier transform of (see Lee, 2004, Eq. (5.1)):

$$B_t(K, T - t) = p(t, T) \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-ik\omega - \gamma k} e^{s_t(\gamma + i\omega)} M_{X_{T|t}}(\gamma + i\omega)}{\gamma + i\omega} d\omega. \quad (5)$$

The technical requirement is weaker; it is sufficient that $E^{\mathbb{Q}}[S_T^\gamma] < \infty$.

2.2.1. The Gauss–Laguerre method

The valuation formula is modified once more before describing the numerical integration method. Since any price is a real number we have that

$$\begin{aligned} C_t(K, T - t) &= p(t, T) \frac{1}{2\pi} \int_{-\infty}^{\infty} \Re \left(\frac{e^{-ik\omega - \gamma k} e^{s_t(1+\gamma+i\omega)} M_{X_{T|t}}(1+\gamma+i\omega)}{(\gamma+i\omega)(1+\gamma+i\omega)} \right) d\omega, \\ &= p(t, T) \frac{1}{\pi} \int_0^{\infty} \Re \left(\frac{e^{-ik\omega - \gamma k} e^{s_t(1+\gamma+i\omega)} M_{X_{T|t}}(1+\gamma+i\omega)}{(\gamma+i\omega)(1+\gamma+i\omega)} \right) d\omega, \end{aligned}$$

where the last equality comes from the fact that the real part of the integrand is an even function of ω . This integral can now be approximated by a Gauss–Laguerre quadrature formula. In general we have that the Gauss–Laguerre quadrature formula approximates an exponentially weighted integral from zero to infinity as

$$\int_0^{\infty} e^{-x} f(x) dx \approx \sum_{j=1}^n w_j^{(n)} f(x_j^{(n)})$$

(see e.g. Abramowitz and Stegun, 1972, p. 890). This paper use $n = 100$, which has turned out to be sufficient to get reasonably accurate prices compared to the accuracy of the observed market prices. We do not claim that this holds true in general, only for the models and options considered. Some care should be taken for short maturities and deep in the money options. The weights and abscissas are pre-calculated and stored in order to increase the computational efficiency of the algorithm. The complexity of the Gauss–Laguerre pricing algorithm will then be of the order $n \times n_{\text{option}}$, i.e. the number of weights times the number of different options.

3. Calibration

Robust calibration of option valuation models to real world option prices is non-trivial but important for good performance. Calibration is entangled with several difficulties, rendering many standard estimators suboptimal. This section will review standard calibration methods, discuss some of the limitations of these and suggest a new framework leading to better use of data.

A serious data problem is the absence of a single quoted market price. Instead, ask prices $c_t^{\text{Ask}}(K_i, \tau_i)$ and bid prices $c_t^{\text{Bid}}(K_i, \tau_i)$ are quoted. The market price is often *approximated* by the mid-price defined as

$$c_t^*(K_i, \tau_i) = \frac{c_t^{\text{Ask}}(K_i, \tau_i) + c_t^{\text{Bid}}(K_i, \tau_i)}{2} \quad (6)$$

and used when calibrating the option valuation models. This may not be an ideal approximation as any parameter vector yielding option values inside the bid–ask spread is acceptable with respect to the arbitrage conditions. Using the mid-price often leads to overfitting of data, and increases the risk for bad hedges.

Other problems can be attributed to parametric modelling. The vast majority of all option valuation models are based on a parametric model for the underlying asset. The postulated stochastic process modelling the underlying asset will only be an approximation of the real world price process, and will be unable to capture finer details of data. The properties of the real price process will also change over time.

A common remedy is to use a simple model that is being recalibrated frequently. Simple models are less prone to overfitting, and it can be argued that traders, knowing the limitations of simple models, can correct most of their pitfalls. The latter claim is valid for transparent (simple) models, but not necessarily for advanced models.

3.1. Least squares methods

The dominating (textbook) calibration method, cf. Bates (1996), Hull (2002), Cont and Tankov (2004) and Schoutens et al. (2004) is weighted least squares (WLS), i.e. taking the parameter (vector) that minimize the weighted sum of the

squared difference between the observed mid-price and the model predicted price

$$L_t^{\text{WLS}}(\theta) = \sum_{s=1}^t \sum_{i=1}^{N_s} \lambda_{s,i} (c_s^*(K_i, \tau_i) - c_s^{\text{Model}}(K_i, \tau_i; \theta))^2, \quad (7)$$

$$\hat{\theta}_t = \arg \min_{\theta \in \Theta} L_t^{\text{WLS}}(\theta). \quad (8)$$

It is statistically optimal (minimal variance of the estimates) to choose $\lambda_{s,i}$ as the inverse of the variance of the residuals, although economic argument may suggest other weights. It is common to choose $\lambda_{s,i} = 0 \ \forall s < t$ to increase the adaptiveness of the calibration, and to choose $\lambda_{t,i}$ as constant or proportional to the inverse of squared bid–ask spread, thereby relating the size of the ask–bid spread to the quality of the quoted prices

$$\lambda_{t,i} \propto \frac{1}{(c_t^{\text{Ask}}(K_i, \tau_i) - c_t^{\text{Bid}}(K_i, \tau_i))^2}. \quad (9)$$

The WLS estimator using these weights has one strong limitation. An implicit assumption is made that old data do not improve the estimation, as the summation is restricted to current data. The implication of this is quite dramatic, as it implies that the past is of limited use to predict the current prices, and hence currently available information are of limited use for predictions of the future! Another problem with this calibration setup is that it uses a small set of observations, easily overfitting data. It is also known to be numerically difficult to find the parameters minimizing the loss function. The latter is discussed in e.g. [Cont and Tankov \(2004\)](#), where a penalty function $P(\theta, \theta_0)$ is added to the loss function

$$L_t^{\text{PWLS}}(\theta) = L_t^{\text{WLS}}(\theta) + \zeta P(\theta, \theta_0). \quad (10)$$

Here, θ_0 is a reference parameter vector and $P(\theta, \theta_0)$ is a positive, locally convex function satisfying $P(\theta, \theta_0) \geq 0$ and $P(\theta_0, \theta_0) = 0$. Assuming that $P(\theta, \theta_0)$ is globally convex and that $L_t^{\text{WLS}}(\theta)$ is well defined will, for a sufficiently large ζ make $L_t^{\text{PWLS}}(\theta)$ globally convex, thus improving the performance of most numerical optimization schemes. However, questions regarding exclusion of data remain unanswered. The penalty may also influence the loss function too much, leading to biased parameters.

Several penalties have been used in the literature. Relative entropy is discussed in [Cont and Tankov \(2004\)](#). Another approach related ridge regression or Tikhonov regularization (see e.g. [Draper and Smith, 1998, Chapter 17](#)) is a quadratic penalty. We use the latter, using the deviation from the previously estimated parameters ($\theta_0 = \theta_{t-1}$) as penalty. Hence, we get

$$P(\theta, \theta_{t-1}) = (\theta - \theta_{t-1})^2. \quad (11)$$

This penalty decrease the overfitting as the paths of the estimated parameters are smoothed over time, while simultaneously being easy to implement. It is also computationally advantageous, as many numerical optimization routines are exact for quadratic problems. We have used Matlabs routine for non-linear least squares minimization, `lsqnonlin`, to obtain the parameters and latent state.

The penalty parameter ζ can be chosen using different strategies. The literature on ridge regression suggests that the parameter is chosen using the connections to Bayesian statistics, cross validation, Morozov discrepancy principle or L-curve methods. We use the discrepancy principle discussed in [Cont and Tankov \(2004\)](#), choosing ζ as

$$\hat{\zeta} = \sup \left\{ \zeta : \sum_{i=1}^{N_t} \lambda_{t,i} (c_t^*(K_i, \tau_i) - c_t^{\text{Model}}(K_i, \tau_i; \theta(\zeta)))^2 \leq d_0 d_1 \right\}, \quad (12)$$

$$d_0 = \sum_{i=1}^{N_t} \lambda_{t,i} (c_t^*(K_i, \tau_i) - c_t^{\text{Model}}(K_i, \tau_i; \theta_t^{\text{WLS}}))^2, \quad (13)$$

where d_1 is chosen as $d_1 = 1.5$. This improves the stability of the optimization, cf. [Cont and Tankov \(2004\)](#).

3.2. Partially observed models

The least squares methods presented above do not explicitly model the time variation of the parameters, and fails to separate the effects from the measurement errors and the time varying parameters consistently. Furthermore, the least squares estimators do not use the structure of the latent processes when estimating the current value. This is hardly optimal use of data.

Using the state-space formulation of the model, and interpreting the market price ambiguity generated by the ask–bid spread as measurement errors addresses some of the problem mentioned. The model is then given as

$$c_t^*(K_i, \tau_i) = c_t^{\text{Model}}(K_i, \tau_i; \theta, V_t) + \varepsilon_t, \quad (14)$$

$$V_t \sim p(V_t | V_{t-1}; \theta), \quad (15)$$

where ε_t is a zero mean random vector having covariance matrix $\mathbf{R} = \text{diag}(\lambda_{t,i})/16$, $\lambda_{t,i}$ is defined as in Section 3.1, θ is a constant parameter vector and V_t is the latent factor (volatility) propagated using the exact transition kernel. Filter algorithms use all observations (not only the most recent observation) to estimate the latent process, and will weight them optimally. This setup can be found in e.g. [Bakshi et al. \(2007\)](#), where an unscented Kalman filter was used to estimate the latent volatility and the parameters were estimated offline using the quasi-maximum likelihood estimator.

The parameters can also be estimated by augmenting the latent state V_t with the parameter vector θ , thereby estimating the parameters online using non-linear filtering. This is computationally easier than maximizing the quasi-likelihood. The online calibration model is given as

$$c_t^*(K_i, \tau_i) = c_t^{\text{Model}}(K_i, \tau_i; \theta_t, V_t) + \varepsilon_t, \quad (16)$$

$$\theta_t = \theta_{t-1}, \quad (17)$$

$$V_t \sim p(V_t | V_{t-1}; \theta_{t-1}). \quad (18)$$

This is a dynamic Bayesian method, but does not address the problem of time varying parameters. Adaptive estimation is achieved if the state-space model is reformulated as a self-organizing state-space model, augmenting the latent state vector V_t with (a transformed version of) the parameter vector $\tilde{\theta}_t = \alpha(\theta_t)$, where the transformed parameter vector is modelled as a random walk, cf. [Anderson and Moore \(1979\)](#) and [Kitigawa \(1998\)](#). This defines the self-organizing state-space model as

$$c_t^*(K_i, \tau_i) = c_t^{\text{Model}}(K_i, \tau_i; \tilde{\theta}_t, V_t) + \varepsilon_t, \quad (19)$$

$$\tilde{\theta}_t = \tilde{\theta}_{t-1} + \eta_t^\theta, \quad (20)$$

$$V_t \sim p(V_t | V_{t-1}; \tilde{\theta}_{t-1}). \quad (21)$$

Adaptively estimating the parameters and the latent volatility is thus transformed into filtering of latent processes, a problem for which statistical methods have been developed during the last decades, [Anderson and Moore \(1979\)](#), [Doucet \(2001\)](#) and [Künsch \(2001\)](#).

3.2.1. Non-linear filtering

Finding the distribution of a latent, vector valued Markov process from noisy measurements is called the filtering problem. Let $\{y_{1:n}\}$ be a shorthand notation for the observations $y_1, \dots, y_n$ and let $\{x_n\}$ be the value of the latent process at time n . Furthermore, assume the measurement density $p(y_n | x_n)$, the state transition density $p(x_{n+1} | x_n)$ and initial state $p(x_0)$ to be known. These components can then be combined through Bayes' formula giving filter density $p(x_n | y_{1:n})$. An attractive feature is that the filter density can be computed recursively, although the resulting filter equations can rarely be solved in closed form. Two exceptions that can be solved are the finite state filter (the hidden Markov model filter) and the Kalman filter, which solves the filter equations for linear, Gaussian systems, see [Kalman \(1960\)](#) and [Kalman and Bucy \(1961\)](#).

The extended Kalman filter (EKF), see [Anderson and Moore \(1979\)](#) and [Maybeck \(1982\)](#), solve the non-linear filtering problem by approximating the transition kernels by Gaussian distributions, leading to explicit filter equations. The EKF is obtained by local linearization of the state transition process and the measurement equation. Assume that

the system of interest can be written as

$$y_t = h(u_t, x_t) + \varepsilon_t, \quad (22)$$

$$x_t = f(x_{t-1}, \eta_t), \quad (23)$$

where $\varepsilon \sim N(0, \mathbf{R})$, $\eta \sim N(0, \mathbf{Q})$ and u_t is the set of $\mathcal{F}_t$ measurable variables (e.g. $S_t, K_i, \tau_i, \dots$). The EKF is then given by

Extended Kalman Filter

Initialization: Let $p(x_0) \sim N(m_0, P_0)$.

Propagate: Propagate mean and covariance as

$$m_{t+1|t} = f(m_t, 0), \quad (24)$$

$$P_{t+1|t} = F(m_t, 0)P_{t|t}F^T(m_t, 0) + G(m_t, 0)\mathbf{Q}G^T(m_t, 0). \quad (25)$$

Updating: The mean and covariance of the filter density are given by

$$K = P_{t+1|t}H^T(m_{t+1|t})(H(m_{t+1|t})P_{t+1|t}H^T(m_{t+1|t}) + \mathbf{R})^{-1}, \quad (26)$$

$$m_{t+1|t+1} = m_{t+1|t} + K(y_{t+1} - h(u_{t+1}, m_{t+1|t})), \quad (27)$$

$$P_{t+1|t+1} = P_{t+1|t} - KH^T(m_{t+1|t})P_{t+1|t}, \quad (28)$$

and repeat the propagation and updating steps until $t = T$, where

$$F(m, \eta) = \frac{\partial f}{\partial x}(m, \eta), \quad G(m, \eta) = \frac{\partial f}{\partial \eta}(m, \eta), \quad H(m) = \frac{\partial h}{\partial x}(u_{t+1}, m). \quad (29)$$

The linearization requires the system to be fairly linear to be accurate, and the filter may diverge if the system is too non-linear.

A simple improvement of the EKF is the iterated EKF (IEKF), see [Maybeck \(1982\)](#), in which the update recursion is replaced by iterating the following recursions a fixed number of iterations or until convergence.

Iterated Extended Kalman Filter

Updating (modified): Iterate until convergence

$$K_{i+1} = P_{t+1|t}H^T(z_i)(H(z_i)P_{t+1|t}H^T(z_i) + \mathbf{R})^{-1}, \quad (30)$$

$$z_{i+1} = m_{t+1|t} + K_{i+1}(y_{t+1} - h(u_{t+1}, z_i) - H^T(z_i)(m_{t+1|t} - z_i)), \quad (31)$$

$$P_{t+1|t+1} = P_{t+1|t} - K_{i+1}H^T(z_{i+1})P_{t+1|t}, \quad (32)$$

starting the recursion using $z_0 = m_{t+1|t}$.

The IEKF changes the point of linearization from $m_{t+1|t}$ to z , producing more accurate approximations of the filter distribution when the measurements are informative, see [Lefebvre et al. \(2004\)](#). In fact, the linearization errors in the IEKF update are smaller than the linearization errors in EKF and unscented Kalman filter updates. This makes the IEKF our filter of choice as the latent dynamics in all models in this paper are linear.

Other non-linear filters can also be used, e.g. [Lindström et al. \(2006\)](#) implements both a standard bootstrap filter and a hybrid bootstrap filter but does not find any support that these are competitive when using a modest number of particles compared to the IEKF. This indicates that the measurement equation is sufficiently linear for the data used in this paper.

3.2.2. Implementation

The filter algorithm cannot be applied to data without specifying an initial distribution of the latent states. We have initialized the filter using the asymptotics given by the WLS estimator, i.e. taking $m_0 = \theta_1^{\text{WLS}}$ and $P_0 = \text{Cov}(\theta_1^{\text{WLS}})$. This is a reasonable prior and ensures that the filter has good starting values, and is not worse off than the WLS when comparing the calibration methods.

Many parameters and latent states are restrained to compact or semi-infinite spaces. These have been transformed onto the real line, as the filters are based on Gaussian approximations. We denote the transformed parameters by $\hat{\theta}$,

cf. Eqs. (19)–(21). Neglecting to transform the parameters may cause the algorithm to break down. Standard, bijective transformations have been used, e.g. taking the logarithm of variables with support on the positive real line and applying the inverse hyperbolic tangent (arctanh) to variables with support on $[-1, 1]$. The transition kernel for latent volatility is approximated by a Gaussian distribution, using the exact conditional moments. The implemented model is given by

$$c_t^*(K_i, \tau_i) = c_t^{\text{Model}}(K_i, \tau_i; \tilde{\theta}_t, V_t) + \varepsilon_t, \quad (33)$$

$$\tilde{\theta}_t = \tilde{\theta}_{t-1} + \eta_t^\theta, \quad (34)$$

$$V_t = \xi^{\mathbb{P}}(1 - e^{-\kappa^{\mathbb{P}} \Delta}) + e^{-\kappa^{\mathbb{P}} \Delta} V_{t-\Delta} + \sqrt{P_t} \eta_t^V, \quad (35)$$

$$P_t = \frac{\sigma_V^2 \xi^{\mathbb{P}}}{2\kappa^{\mathbb{P}}} + \frac{\sigma_V^2 (V_{t-\Delta} - \xi^{\mathbb{P}})}{\kappa^{\mathbb{P}}} e^{-\kappa^{\mathbb{P}} \Delta} + \frac{\sigma_V^2 (\xi^{\mathbb{P}} - 2V_{t-\Delta})}{2\kappa^{\mathbb{P}}} e^{-2\kappa^{\mathbb{P}} \Delta}. \quad (36)$$

The latent volatility evolves according to the $\mathbb{P}$ -dynamics, while our parameters are $\mathbb{Q}$ -parameters. These are related through risk premiums λ_V according to

$$\kappa = \kappa^{\mathbb{P}}(1 + \lambda_V), \quad \xi = \frac{\xi^{\mathbb{P}}}{1 + \lambda_V}. \quad (37)$$

Several studies estimating the risk premiums have found that their estimates are either weakly significant or even insignificant, cf. [Polson and Stroud \(2003\)](#) on S&P 500 data and [Carr and Wu \(2007\)](#) on currency option data. We have been unable to get statistically significant estimates on our (shorter) set of data, and have therefore approximated the $\mathbb{P}$ -dynamics with the $\mathbb{Q}$ -dynamics ($\lambda_V = 0$), although simultaneous estimation of $\mathbb{Q}$ -parameters and risk premiums is preferred when possible.

What remains is to specify the covariance matrix of η^θ . We have for computational and statistical simplicity assumed that increments of each parameter are pair wise independent. It is well known that stochastic volatility is clearly needed to fit data, while the gain from introducing jumps is smaller (but still significant). We have therefore chosen to allow for different covariances, assigning covariance Q^c to parameters related to the continuous (stochastic volatility) component and covariance Q^j to parameters related to the jump component. These are different, while covariances in each group are identical. The size of the covariances has been estimated using quasi-maximum likelihood on an independent set of data, simulated using parameters that resemble those we are interested in estimating.

4. Simulation study

The methods are first tested in a simulation study. Data have been simulated from the Bates model, cf. Section 2.1, as this model includes jumps and stochastic volatility. The initial $\mathbb{Q}$ -parameters were chosen according to [Bakshi et al. \(1997\)](#) to mimic S&P 500 index as closely as possible, see [Table 1](#).

The Bates model has seven parameters and one latent state which are grouped together in the augmented latent state vector, $x = (\kappa, \xi, \sigma_V, \rho, V_t, \lambda, \mu, \delta)$. Prices have been generated with constant parameters except V_t which follows a stochastic process according to (2), and ξ and ρ which have been varied as depicted in [Fig. 1](#). The purpose of this setup is to test whether the methods can track changing parameters.

The index level and latent volatility were simulated under the $\mathbb{P}$ -dynamics, where we used $\lambda_V = -0.25$, while the jump parameters were identical under $\mathbb{P}$ and $\mathbb{Q}$. Using different volatility parameters under $\mathbb{P}$ and $\mathbb{Q}$, while filtering using only $\mathbb{Q}$ -parameters will emulate real data, and test our approximation (cf. Section 3.2) under controlled conditions.

The total length of the data set is 200 trading days. Hundred European call options over a range of 20 strikes with moneyness between 0.65 and 1.35 and five maturities were generated each trading day. Our options data have bid–ask

Table 1
Parameters for the Bates model used in the simulation study

Parameter	S_0	κ	ξ	σ_V	ρ	V_0	λ	μ	δ
Value	1000	2.03	0.04	0.38	−0.7	0.0576	0.59	−0.05	0.07

Fig. 1. Bates simulated data. True parameters (solid lines), WLS estimates (circles) and IEKF estimates (dots). ξ and ρ are changing over time.

spread of approximately \$1, so we have added noise to the option prices, cf. Eq. (19), to mimic the mid-price uncertainty present in the real world option prices. The noise is Gaussian with a standard deviation of one-fourth of the spread, cf. $\mathbf{R} = \text{cov}(\varepsilon) = \frac{1}{16}$. The motivation behind this choice is that we expect at least 95% of the true option prices to be within the spreads. Interpreting the error as a pure truncation error would give a slightly different $\mathbf{R}$. The set of option prices was split into an estimation set of 40 options per day and a validation set of 60 options per day. The options in the different sets were selected at random each day.

Calibration can be seen as a mapping between, in our case, European calls and a set of latent processes. If the validation is performed using the same mapping, e.g. by pricing out-of-sample calls, the interpolation capacity of the calibrated pricing model is tested. If on the other hand the quality of the calibrated model is to be tested, a different mapping is more suitable. A convenient choice in this simulation study is the European binary option, which can be priced in the same framework as the European call as seen in Section 2.2. Prices of binary options are thus calculated both for the true parameters and for the filtered states and compared using RMSE.

The time consumption has been measured by the average time consumption per step, i.e. the average time consumption of daily recalibration. Since the time consumption is dependent of factors not specific to the actual estimation method, e.g. the performance of the computer used, we are more interested in the ordering among the methods than the actual numbers.

The parameters have been estimated using two different filters: EKF and IEKF. The state noise covariance matrix $\mathbf{Q}$ was determined using quasi-maximum likelihood estimation as discussed in Section 3.2. We found that $\mathbf{Q}^c = 5 \times 10^{-3}$ and $\mathbf{Q}^j = 10^{-4}$ to be optimal.

As a benchmark the parameters have also been estimated using standard and penalized WLS (PWLS), cf. Section 3.1. The least squares estimators were initialized with the true parameter values. The PWLS used $d_1 = 1.5$, giving $\hat{\zeta} \approx 10$. We have used $\hat{\zeta} = 10$ throughout the simulation study.

4.1. Results

The IEKF and WLS estimated parameters are depicted in Fig. 1 in addition to the true parameter values. It can be seen that the IEKF estimates are closer to the true parameter values than the WLS estimates. We have not plotted the

Table 2
Summary statistics for the simulated Bates model

Filter	Set	IS (%)	RMSE calls	RMSE parameters	RMSE binary (10^{-5})	Average time (s)
EKF	Est.	91.50	0.0924	0.0017	0.0572	0.5933
	Val.	88.63	0.1078	–	–	–
IEKF	Est.	96.86	0.0546	0.0018	0.0354	2.8916
	Val.	93.65	0.0712	–	–	–
PWLS	Est.	96.98	0.0536	0.0066	0.0556	6.4099
	Val.	93.33	0.0732	–	–	–
WLS	Est.	97.04	0.0521	2.8305	0.1452	135.67
	Val.	92.43	0.0783	–	–	–

The RMSE for parameters are over all eight parameters. Bold face indicates the best method for that measure.

EKF and PWLS, these being intermediary in performance. WLS estimates far from the true values were moved to the edge the subplots. This problem was mostly present in the estimates of the jump intensity λ and jump volatility δ as can be seen in the figure. The IEKF (and EKF, PWLS) estimates are more stable over time compared to the WLS estimates, and all the methods, possibly except WLS, estimate the latent process V_t very well. Even better, all methods are capable of tracking time varying parameters with reasonable accuracy. This holds both for slowly varying parameters ξ and ρ and for discontinuities in the ξ parameter.

Table 2 displays the numerical results of the measured fit for the different estimators. As expected (defined) the WLS have the best fit in sample. On the validation set the IEKF and the PWLS perform the best. The PWLS drops more in performance on the validation set, indicating a slight overfitting, but still much better than the WLS. When measuring the RMSE of the call prices for the different methods we again see the WLS performing best in sample and the IEKF out of sample.

In terms of parameter RMSE, the filter methods clearly outperform the least squares methods, the WLS being several orders of magnitude larger. This result is consistent with the overfitting we saw on the European call options and the graphical results in Fig. 1. When considering European binary options, we again see a clear advantage for the IEKF, while the EKF and PWLS being approximately equal. The RMSE for the WLS is much larger than for any of the other methods.

In terms of speed, the EKF is the clear winner. It is our subjective belief from simulations using the Heston model and Bates model that the WLS estimator suffers from the curse of dimensionality more than the filter methods. The larger parameter space requires more function evaluation for the minimization routine, yielding increased computational time or decreased precision. The estimation set in our simulation is quite large, and it is not unexpected that the interpolation capacity is sufficiently tuned on the estimation set to also perform well on the validation set for the WLS. If the number of options per day were smaller or varied, the WLS would be more prone to overfitting.

It can be observed that the filter does not seem to suffer badly (compared to the WLS and PWLS) from using only $\mathbb{Q}$ -parameters instead of using $\mathbb{Q}$ - and $\mathbb{P}$ -parameters simultaneously. We are therefore using the same approximation ($\lambda_V = 0$) on real data.

5. Empirical study

We have used daily data on S&P 500 index options, from November 5th, 2001 to May 5th, 2003, thus excluding data containing a short term effect following September 11th. The original data set consisted of 38 225 quotes (date, strike, ask price, bid price, index level, time to maturity and risk free interest rate). The raw data were processed according to the following rules:

1. All options having time to maturity less than 6 trading days were excluded. These options are illiquid and highly sensitive to changes in any variable.
2. Options having bid–ask spread larger than 5 have been excluded, in order to eliminate illiquid options. Also options having zero or negative ≤ 0 bid–ask spread have been eliminated as these are an effect of non-synchronous trading.

3. The least liquid option when call and put options having identical strike and time to maturity is available has been eliminated. It is well known that in-the-money options are traded infrequently compared to at-the-money or out-of-the-money options. We have therefore excluded in-the-money options when at-the-money or out-of-the-money options are available.

A total of 24 930 unique quotes remain after the processing. It is convenient to transform all quoted put options into call options, using the put–call parity.

In practice r , the continuously compounded short rate (cf. Eq. (1)) is not constant and known. If we want to price a claim at time t with maturity at time T , it is possible to replace r with the zero-coupon-bond (ZCB) yield over the corresponding time period up to maturity (t, T) and replace $\mathbb{Q}$ with the corresponding forward measure, $\mathbb{F}^T$ instead. A drawback with this approach is that we get different dynamics for the underlying stock for different maturities, i.e. for each combination of t and T . This is of course due the fact that the ZCB yield varies with t and T . From a valuation perspective this causes no serious problem, simply replacing r with the corresponding zero-coupon-yield and set $p(t, T)$ equal to the ZCB price. The pricing formula Eq. (4) will look exactly the same.

Another adjustment is needed to account for dividends. It is common to approximate the dividend by a continuously compounded dividend yield that depends on time and on time to maturity, hereafter denoted $q(t, t + \tau)$. Ignoring the dividend yield causes the transformed option prices to be discontinuous in the strike level. The dividend yield is estimated such that the option prices are continuous in the strike level.

It is shown in e.g. Hull (2002) that dividends can be modelled by replacing the current value of the underlying equity or index S_t by $S_t e^{-q(t, t+\tau)\tau}$, when modelling a European call option maturing at $t + \tau$. However, $q(t, t + \tau)$ is not quoted on the market, why we derive it from option quotes. It is well known from the put–call parity that

$$c_t(K_i, \tau) - p_t(K_i, \tau) = S_t e^{-q(t, t+\tau)\tau} - K_i p(t, t + \tau). \quad (38)$$

It is possible to estimate the corrected index level $S_t e^{-q(t, t+\tau)\tau}$. Rearranging Eq. (38) and adding a noise term e_i gives

$$S_t e^{-q(t, t+\tau)\tau} = c_{K_i, \tau}(t) - p_{K_i, \tau}(t) + K_i p(t, t + \tau) + e_i. \quad (39)$$

The corrected index level can now be estimated by the sample mean, taken over overlapping mid-prices. The estimate is stable over different K_i and reasonably stable over time.

5.1. Results

We have estimated three models on market data from options on the S&P 500 index, namely Heston, Bates and NIG-CIR. Eighty percent of the processed data was used for estimation and the remainder for validation. The IEKF, PWLS and WLS parameter estimates are presented in Figs. 2, 3 and 4. As in the simulation study, the axes are truncated and some WLS estimates that obtained very large (absolute) values have been moved to the upper or lower edge of the plots.

For the Heston model, being a sub-model of the Bates model, we used the state covariance found in the simulation study. For the PWLS we set ζ as given by Eq. (12) to 10, also inherited from the simulation study. A posterior numerical calculation proved this a sensible choice. The result from the Heston estimation is depicted in Fig. 2. The WLS and IEKF estimates (and PWLS) agree on ξ , ρ and V_t . However, the WLS is less robust on κ and σ_V giving highly varying estimates.

The Bates model was used in the simulation study on a set of data designed to be similar to the actual S&P 500 data. We optimized with respect to Q and ζ in the simulation and used the same values for the market data set. Also, ζ was checked posterior to have a reasonably correct value. The result from the Bates estimation is depicted in Fig. 3. For the highly influential latent volatility process V_t , the WLS and IEKF agree, while on the other parameters, the WLS is much more erratic, cf. the simulation study in Section 4. The PWLS and IEKF on the other hand, agree to a large extent. It should be noted though that all parameters move a lot and are from being constant. For the jump intensity λ the WLS sometimes gave extreme estimates (this has been truncated in the figure), something it counteracts by lowering both the expected value μ and standard deviation δ . Also note the similarity with the Heston model of the volatility process and its related parameters for the IEKF and PWLS. This, however, does not hold for the WLS.

The NIG-CIR model proved to be somewhat harder to find good parameter settings since we had no prior experience from simulation study. We expect similar variation of the parameters in the NIG-CIR model as in the Bates model, and

Fig. 2. Heston on market data. WLS (circles), PWLS (thin line) and IEKF estimates (thick line).

Fig. 3. Bates on market data. WLS (circles), PWLS (thin line) and IEKF estimates (thick line).

Fig. 4. NIG-CIR on market data. WLS (circles), PWLS (thin line) and IEKF estimates (thick line).

Table 3

The Heston, Bates and NIG-CIR models have been calibrated to a number of S&P 500 data sets using the different estimation methods

Filter	Set	Heston		Bates		NIG-CIR	
		IS (%)	RMSE	IS (%)	RMSE	IS (%)	RMSE
EKF	Est.	71.72	0.8080	76.68	0.6252	75.00	0.9331
	Val.	70.27	0.7772	76.20	0.6335	74.13	0.9651
IEKF	Est.	77.08	0.4609	85.99	0.3615	87.93	0.3242
	Val.	76.08	0.4781	85.13	0.3615	86.93	0.3309
PWLS	Est.	75.04	0.4796	83.35	0.3796	86.48	0.3450
	Val.	74.25	0.4934	82.69	0.3830	85.49	0.3466
WLS	Est.	77.75	0.4569	89.57	0.2914	87.59	0.3266
	Val.	76.06	0.4760	82.69	0.3201	85.78	0.3393

The table contains the global fit error measurements for the calibrated model in sample (estimation set) and out of sample (validation set) for the different estimation methods.

used the same settings. Other parameter values were also tested but no improvement was achieved. The result from the NIG-CIR estimation is depicted in Fig. 4. Note that for the latter part of the data set, the WLS estimator has given fairly unrealistic estimates, probably getting stuck in a local minimum. This is also seen in Table 3. Further, the ratio α/β , having a similar interpretation as the correlation in the Bates model, is fairly constant for the IEKF and PWLS, but varies a lot for the WLS.

Global fit of the models has been measured in terms of inside spread (IS) and RMSE of the predicted prices in the same way as in the simulation study. These results are presented in Table 3. No evaluation with respect to more exotic contracts has been made.

Since the validation set only consists of approximately 20% of the options, we expect the interpolation capacity from the WLS estimates to be sufficiently tuned to also perform well out of sample, something that is confirmed in Table 3. This is most evident in the simpler Heston model, while for the more complex models, evidence of overfitting is seen. On the other hand we expect the IEKF to perform up to par with the WLS whilst still achieving stable parameter estimates. This is confirmed by the parameter trajectories shown in Figs. 2–4 as well as the statistics in Table 3. As in the simulation study, the IEKF performs best out of sample on all three models, indicating a well calibrated market model and not just interpolation of the prices in the estimation set. It is worth noting that the WLS still overfits, as is most evident in the Bates model where there is a substantial difference between in sample and out of sample IS.

6. Conclusions

The filter technique proposed in the paper has been evaluated on simulated and market option data. The filter estimates have been compared to the standard WLS estimates as well as to estimates from a PWLS method. The overall impression is that filter methods provide faster, more stable and more robust estimation than WLS.

When dealing with market quoted prices, the use of mid-price introduces a risk for overfitting. We find abundant evidence of overfitting in the WLS estimates. The limitation of the standard WLS implementation is that it only uses data from the current observation, whereas the filter methods make optimal use of past and current observations. An intermediary approach to reduce overfitting is the PWLS, but this lacks the optimal weighting of past and current information present in the filter.

We performed a simulation study on the Bates model to evaluate the different methods under controlled circumstances. The WLS achieved a high in sample fit, with the PWLS and IEKF being close behind. The ordering changed out of sample with the IEKF being better than the WLS. This indicates that the WLS overfitted to data. The IEKF on the other hand managed to find the actual parameter values, and therefore excelling on the other metrics, especially when pricing binary options. Even the EKF manages to track the parameters well but the inferior linearization means that it is not capable of fine-tuning the more influential parameters such as the volatility to the extent that the IEKF does. Also, the amount of time required by the different methods makes IEKF stand out even more.

A similar study was performed on S&P 500 option data. The pattern from the simulation study was repeated. The filter methods managed to provide stable estimates across models while the WLS is more erratic and has a poorer performance on the complex models. In fact we experienced some convergence issues for the WLS on the NIG-CIR model. We can only speculate that the parameter estimates from the IEKF and PWLS represent the evolution some actual market model, something that needs to be checked using a different mapping than European calls similar to the binary calls in the simulation study.

These results provide evidence that the self-organizing state-space model, combined with an IEKF is a promising method for option calibration. The IEKF is shown to be both faster and more accurate than standard methods when applied to different market models. We believe that a similar behaviour can be expected from the filter for other models and other data sets.

Acknowledgements

We would like to thank an anonymous referee and the guest editor for helpful comments and suggestions. Erik Lindström gratefully acknowledges financial support from the Bank of Sweden Tercentenary Foundation under Grant P2005-0712, Jonas Ströjby gratefully acknowledges financial support from the Swedish Research Council under Grant 2005-2799 and Magnus Wiktorsson gratefully acknowledges financial support from the Swedish Research Council under Grant 2002-5415.

References

- Abramowitz, M., Stegun, I.A. (Eds.), 1972. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th printing. Dover, New York.
- Anderson, B.D.O., Moore, J.B., 1979. Optimal Filtering. Prentice-Hall, Englewood Cliffs, NJ.
- Bakshi, G., Carr, P., Wu, L., 2007. Stochastic risk premiums: stochastic skewness in currency options, and stochastic discount factors in international economics. *J. Financial Econom.*, in press, doi:10.1016/j.jfineco.2006.12.001.
- Bakshi, G.S., Cao, C., Chen, Z., 1997. Empirical performance of alternative option pricing models. *J. Finance* 52 (5), 2003–2049.

- Barndorff-Nielsen, O.E., 1997. Processes of Normal inverse Gaussian type. *Finance and Stochastics* 2, 41–68.
- Barndorff-Nielsen, O.E., Shepard, N., 2001. Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *J. Roy. Statist. Soc. Ser. B* 63, 167–241.
- Barndorff-Nielsen, O.E., Shepard, N., 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *J. Roy. Statist. Soc. Ser. B* 64, 253–280.
- Bates, D., 1996. Jumps and stochastic volatility: the exchange rate processes implicit in Deutchemark options. *Rev. Financial Stud.* 9, 69–107.
- Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. *J. Political Economy* 81, 637–659.
- Campbell, J.Y., Lo, A.W., MacKinlay, A.C., 1997. *The Econometrics of Financial Markets*. Princeton University Press, Princeton, NJ.
- Carr, P., Madan, D., 1999. Option valuation using the fast Fourier transform. *J. Comput. Finance* 2 (4), 61–73.
- Carr, P., Wu, L., 2003. The finite moment log stable process and option pricing. *J. Finance* 58, 753–778.
- Carr, P., Wu, L., 2004. Time-changed Lévy processes and option pricing. *J. Financial Econom.* 71, 113–141.
- Carr, P., Wu, L., 2007. Stochastic skew in currency options. *J. Financial Econom.*, in press, doi:10.1016/j.jfineco.2006.03.010.
- Carr, P., Geman, H., Madan, D., Yor, M., 2002. The fine structure of asset returns: an empirical investigation. *J. Business* 75, 305–332.
- Carr, P., Geman, H., Madan, D., Wu, L., Yor, M., 2003a. Option pricing using integral transforms. A slideshow presentation available at: (<http://www.math.nyu.edu/research/carrp/papers/pdf/integtransform.pdf>).
- Carr, P., Geman, H., Madan, D., Yor, M., 2003b. Stochastic volatility for Lévy processes. *Math. Finance* 13, 345–382.
- Chernov, M., Gallant, A.R., Ghysels, E., Tauchen, G., 2004. Alternative models for stock price dynamics. *J. Econometrics* 116, 225–257.
- Cont, R., Tankov, P., 2004. *Financial Modelling with Jump Processes*. Chapman & Hall, London.
- Derman, E., Kani, I., 1994. Riding on a smile. *Risk* 7 (2), 32–39.
- Doucet, A., de Freitas, N., Gordon, N. (Eds.), 2001. *Sequential Monte Carlo Methods in Practice*. Springer, New York.
- Draper, N.R., Smith, H., 1998. *Applied Regression Analysis*. third ed.. Wiley, New York.
- Duffie, D., Pan, J., Singleton, K., 2000. Transform analysis and asset pricing for affine jump diffusions. *Econometrica* 68, 1343–1376.
- Dupire, B., 1994. Pricing with a smile. *Risk* 7 (1), 18–20.
- Heston, S.L., 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financial Stud.* 6, 327–343.
- Huang, J.-Z., Wu, L., 2004. Specification analysis of option pricing models based on time-changed Lévy Processes. *J. Finance* 59, 1405–1440.
- Hull, J.C., 2002. *Options, Futures, and Other Derivative Securities*. fifth ed. Prentice-Hall, Englewood Cliffs, NJ.
- Hull, J.C., White, A., 1987. The pricing of options on assets with stochastic volatility. *J. Finance* 42 (2), 281–300.
- Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. *Trans. ASME Ser. D J. Basic Eng.* 82, 35–45.
- Kalman, R.E., Bucy, R.S., 1961. New results in linear filtering and prediction. *Trans. ASME Ser. D J. Basic Eng.* 83, 95–107.
- Kitigawa, G., 1998. A self-organizing state-space-model. *J. Amer. Statist. Assoc.* 93 (443), 1203–1215.
- Künsch, H.R., 2001. State space and hidden Markov models. In: Barndorff-Nielsen, O.E., Cox, D.R., Klüppelberg, C. (Eds.), *Complex Stochastic Systems*, Monographs on Statistics and Applied Probability, vol. 87. Chapman & Hall, New York, pp. 109–173.
- Lee, R.W., 2004. Option pricing by transform methods: extensions, unification, and error control. *J. Comput. Finance* 7, 51–86, an augmented version is available at: (<http://www.math.uchicago.edu/~rl/dft.pdf>).
- Lefebvre, T., Bruyninckx, H., De Schutter, J., 2004. Kalman filters for non-linear systems: a comparison of performance. *Internat. J. Control* 77 (7), 639–653.
- Lindström, E., Ströjby, J., Brodén, M., Wiktorsson, M., Holst, J., 2006. Sequential calibration of options. Technical Report 21, Centre for Mathematical Sciences, Lund University.
- Madan, D.B., Seneta, E., 1990. The variance Gamma (VG) model for share market returns. *J. Business* 63, 511–524.
- Maybeck, P., 1982. *Stochastic Models, Estimation and Control*, vols. 1–3. Academic Press, New York.
- Merton, R.C., 1976. Option pricing when the underlying stock returns are discontinuous. *J. Financial Econom.* 5, 125–144.
- Polson, N.G., Stroud, J.R., 2003. Bayesian inference for derivative prices. *Bayesian Statist.* 7, 641–650.
- Schoutens, W., Simons, E., Tistaert, J., 2004. A perfect calibration! Now what? *Wilmott Magazine*, March.