Simultaneous Calibration and Hedging of Options

Jingyi Guo

January 29, 2013

Abstract

This paper implements a variety of different calibration methods applied to the stochastic volatility jump diffusion models and levy models and examines their effects on the performance of standard Δ and Monte Carlo quadratic hedging of vanilla options on the S&P 500 index.

On the methodological side, we derive a sequential algorithm for simultaneous calibration and quadratic hedging of options. In order to perform the various hedging strategies, we have calibrated several stochastic volatility models with/without jumps with the Ensemble Kalman filter. The algorithm is very flexible and can be applied to any model for which we can price options, and simulate paths from.

An empirical examination of the model calibration performance of a stochastic volatility model is provided for options on the S&P 500 index in which the unobservable time varying volatility is jointly estimated with the time varying parameters of the model.

Afterwards, the research focused on testing the efficiency of simultaneous calibration and hedging. A simulation study first applied on model calibration with additional latent state, named spot price process dynamics. This is the foundation of implementation of the method of simultaneous calibration and hedging. Then we make use of the method of simultaneous calibration and hedging on S&P 500 option data with another empirical study. To complete the test of the method, the Delta hedging strategy and two Monte Carlo quadratic hedging strategies were take into account. The hedging issue is of interests not only to academics, but also to actual traders in the option market who often have to hedge their positions in the underlying asset or options market.

For all of examples that we considered, the ensemble Kalman filter worked successfully once a threshold ensemble size was reached.

Acknowledgment

My deepest gratitude goes first and foremost to my supervisor, Dr. Erik Lindström, for his constant encouragement and guidance. He has walked me through all the stages of the writing of this thesis. Without his impressive kindness and patience, I could not have completed this thesis. Without his consistent and illuminating instruction, this thesis could not have reached its present form. His keen and vigorous academic observation enlightens me not only in this thesis but also in my future research.

Second, I would like to express my heartfelt gratitude to Dr. Magnus Wiktorsson, who led me into the world of Statistics.

Last but not least, I would like to thank all the teachers who have taught me and helped me enrich and broaden my knowledge.

Contents

1 Introduction

The three purposes of a model in derivatives pricing are to calibrate, to price and to hedge. Generally, a stochastic dynamical model which includes several controlled parameters, is adopted to estimate the price process of the underlying asset. The prices of commonly traded derivatives are calculated using that model. And the parameters in the model is then calibrated by fitting the predicted prices against available market observed data. Once the model is calibrated so that it is agreement with liquidly traded derivatives, it can be used to price less liquid products of the same type or hedge the derivatives with the underlying asset or/and with other instruments.

Numerous models for parametric option valuation have been proposed in financial literature over several decades after Black & Scholes (Black and Scholes, 1973) worked on option pricing under no arbitrage arguments. These models have included jumps (Merton, 1976), stochastic volatility and local volatility (Heston, 1993) and state dependent diffusion terms (Dupire, 1994; Derman and Kani, 1994).

In general, all these models can be included into two broad groups: complete market models and incomplete market models. Complete markets allow perfect replication and therefore hedging any conceivable payoff structure using a portfolio of traded assets, while such replication is generally impossible in incomplete markets. On the other hand, incomplete market models have richer structure than complete models since they contain more sources of uncertainty.

Research on hedging European style options in these two markets has important practical applications. In particular, market makers set prices so that their net profit from the trade, after deducting hedging costs, has positive expectation. More accurate hedging allows these traders to reduce their bid-ask spread, and thus increase their volume of trade.

In order to limit the hedging error, good parameter estimates which can be estimated from model calibration technique are needed. The risk in such a model calibration is that it can over-fit the data. The most common calibration technique is some version of daily least squares estimation, which provides good in sample predictions but non-robust to outliers (Cont and Tankov, 2004). Lindström et al. (2008) suggest the standard least squares methods should be replaced by a non-linear (Kalman) filter method since the filter methods make optimal use of past and current observations, whereas the standard WLS only uses data from the current observation.

As an alternative to Least square method, the Kalman filter (Kalman et al., 1960) is one of the most well known filter technologies and is an

optimal, minimum mean square error estimator for linear systems. The main reason for Kalman filter's popularity is its optimality and simplicity, as well as, its recursive construction. When the statistical properties of the modeling error and the uncertainty of the observations are assumed to be normally distributed with zero mean and known covariances, Kalman filter gives an unbiased minimum variance estimation for linear stochastic dynamical systems. While system dynamics are instinsically nonlinear, the formalism of the Kalman filter has been extended to nonlinear dynamical systems with the help of linearization around the prior estimation of the state, which is known as Extended Kalman filter (EKF for short), (see Anderson and Moore, 1979). In general, the EKF performs a truncated first order Taylor expansion regarding the current state, to which the linear filter equations are applied, so the traditional Kalman filter can be applied. Unfortunately, the EKF has two important potential drawbacks. First, it suffers divergence problem, due to the deviation of the Jacobian matrices, the linear approximations to the nonlinear functions can be complex causing implementation difficulties. Second, these linearizations can lead to filter instability if the time step intervals are not sufficiently small. To address these problems, Julie and Uhlmann (Julier and Uhlmann, 1997) developed the Unscented Kalman Filter (UKF) to estimate the state of a non-linear system using the unscented transform.The UKF claims a higher accuracy and robustness for non-linear than the EKF in the time series modeling. Instead of linearizing the functions in the EKF, the UKF picks a minimal set of sample points (called sigma points) and propagates this set through the actual non-linear dynamics, These points are chosen such that their mean, covariance and possibly also higher order moments match the Gaussian random variable. The mean and the covariance can be recalculated from the propagated points, yielding more accurate results compared to the ordinary function linearization. A substantial disadvantage of these filters is the prohibitively high cost to store and maintain the error covariance matrix for large-scale problems as well as they approximate all distributions as Gaussian, making them less suitable when calibrating non-Gaussian models.

For this reason, a scheme named the ensemble Kalman filter(EnKF) was developed. It was proposed by Evensen (Evensen, 1994) and refined by Burger (Burgers et al., 1998), has gained popularity among them. In particular, EnKF estimation is widely used in weather forecasting, where the models are of extremely high order and nonlinear, the initial states are highly uncertain, and a large number of measurements are available. However, to our knowledge, it has not been used in derivatives pricing. There exist few textbook discussions of EnKF estimation. A

brief overview of the technique is given in (Daley, 1993) and (Kalnay, 2002). In the weather prediction literature, there exist a large number of papers that make use of the EnKF (Evensen, 1997, 2002).

Contrary to the EKF and UKF, the EnKF constitutes a class of derivative free nonlinear filters. As a Monte-Carlo implementation of the Kalman filter, it represents the distribution of the model state using a collection of state vectors, called "an ensemble", and replaces the covariance matrix by the sample covariance computed from the ensemble. Ensemble members are then forwarded in time by solving the state equations and analyzed by the Kalman filter scheme. Therefore, the EnKF requires no linear approximation of models and no explicit storage of covariance matrix.Both the EnKF and the UKF have their advantages and disadvantages. The good performance of the EnKF has been shown in Hommels et al. (2005).

In the case of the UKF, the sample points are chosen deterministically. In fact, the number of sample points required is of the same order as the dimension of the system, i.e. $2d + 1$, where d is the dimension of states. On the other hand, the number of ensembles required in the EnKF is heuristic. While one would expect that a large ensemble would be needed to obtain useful estimates, the literature on EnKF suggests that an ensemble of size 50 to 100 is often adequate for systems with thousands of states. The accuracy of the state estimates as a function of ensemble size is thus an important research question (Gillijns et al., 2006).

The present paper has three main goals. First, we summarize the steps of the EnKF estimation for model calibration in derivatives pricing. Next, we apply the EnKF to a collection of two dynamic models to obtain inside parameters of these models. In particular, we consider one stochastic volatility jump diffusion model and one Lévy model. After doing this, we apply Monte Carlo quadratic hedging to test its effectiveness. Our goal is to determine the tradeoff between ensemble size and estimation accuracy. Finally, using the results of these studies, we extend the original EnKF to combine EnKF with Monte Carlo hedging strategies for simultaneous calibration and quadratic hedging of options.

The paper is organized as follows. In Section 2, after a brief overview of the four alternative pricing models of interest is given, the theory of option price calculation using Inverse Fourier transformation is discussed and the dynamic hedging strategies used in this paper, namely the Delta hedging and the Monte Carlo quadratic hedging is described. Section 3 expounds steps of Monte Carlo simulation for different models. Section 4 presents the model calibration approaches to estimating the models parameters on the option price data. Section 5 illustrates the method

of simultaneous calibration and hedging technique. A conclusion is summarize in section 6.

2 Theoretical background

2.1 Overview of stock price models

In this paper we consider a large group of the most applied option pricing models in the contemporary literature, i.e.the Black & Scholes (Black and Scholes, 1973) the Heston (Heston, 1993), Bates (Bates, 1996) as well as the NIG-CIR model (Carr and Wun, 2003).

2.1.1 The Black-Scholes model

Some work on modeling stock prices were made before, but it was not until Black and Scholes (1973) work on option pricing under no arbitrage arguments the development was spurred. They used a stock model introduced by Samuelson (1965) which defines the market under risk neutral measure as:

$$\frac{\mathrm{d}S_t}{S_t} = r\mathrm{d}t + \sigma\mathrm{d}W_t,$$

where r is the short risk-free rate, σ is the volatility of asset value and $\mathrm{d}W_t$ is Wiener process.

The differential equation can be solved explicitly for S_t as,

$$S_t = S_0 \exp\left\{(r - \frac{\sigma^2}{2})t + \sigma W_t\right\}.$$

2.1.2 The Heston model

By introducing a hidden, non-negative process expressed by a stochastic differential equation, we are able to model the impact on asset prices of stochastic volatility. Heston (1993) chose to model asset prices with stochastic volatility as follows:

$$\frac{\mathrm{d}S_t}{S_t} = r\mathrm{d}t + \sqrt{V_t}\mathrm{d}W_t^S,$$

where r is the short risk-free rate and the squared volatility is modeled by a mean reverting CIR process:

$$\mathrm{d}V_t = \kappa(\xi - V_t)\mathrm{d}t + \sigma\sqrt{V_t}\mathrm{d}W_t^V. \tag{2.1}$$

where κ is the speed of mean reversion, ξ is the average of volatility, σ is the volatility of volatility.

The diffusion terms $\mathrm{d}W_t^S$ and $\mathrm{d}W_t^V$ are defined as correlated Wiener processes:

$$\mathrm{Cov}(\mathrm{d}W_t^S, \mathrm{d}W_t^V) = \rho \mathrm{d}t.$$

Given the Heston model, the characteristic function of the logarithm of the asset price process, s_t, follow as (see Schoutens et al., 2004) or (see Heston, 1993):

$$\phi_{s_t}(z) = \frac{\exp\left\{\frac{\kappa\xi t(\kappa - \mathrm{i}\rho\sigma z)}{\sigma^2} + \mathrm{i}ztr + \mathrm{i}zs_0\right\}}{\left(\cosh\frac{\gamma t}{2} + \frac{\kappa - \mathrm{i}\rho\sigma z}{\gamma}\sinh\frac{\gamma t}{2}\right)^{\frac{2\kappa\xi}{\sigma^2}}} \exp\left\{-\frac{(z^2 + \mathrm{i}z)V_0}{\gamma\coth\frac{\gamma t}{2} + \kappa - \mathrm{i}\rho\sigma z}\right\}. \tag{2.2}$$

where $\gamma = \sqrt{\sigma^2(z^2 + \mathrm{i}z) + (\kappa - \mathrm{i}\rho\sigma z)^2}$, and s_0 and V_0 are initial values for the log-price process and the volatility process respectively.

2.1.3 The Bates model

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

$$\begin{aligned} \mathrm{d}S_t &= rS_t\mathrm{d}t + \sqrt{V_t}S_t\mathrm{d}W_t^{(S)} + S_t\mathrm{d}J_t, \\ \mathrm{d}V_t &= \kappa(\xi - V_t)\mathrm{d}t + \sigma_V\sqrt{V_t}\mathrm{d}W_t^{(V)}, \end{aligned}$$

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 $\log(1 + \Delta J) \in \mathrm{N}(\mu, \delta^2)$ 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 $\mathbb{Q}$.

Since the jumps are independent of the diffusion part, the characteristic function of the logarithmic asset price proceed under Bates model can be written:

$$\phi_{s_t}(z) = \phi_{s_t}^D(z)\phi_{s_t}^J(z),$$

where the diffusion part of characteristic function is equivalent to that of the Heston model, see Equation (2.2), and the jump part of the characteristic function is:

$$\phi_{s_t}^J(z) = \exp\left\{t\lambda(\mathrm{e}^{\delta^2 z^2/2 + \mathrm{i}(\log(1+k) - \frac{1}{2}\sigma^2)z} - 1)\right\}.$$

2.1.4 The NIG-CIR model

The NIG-CIR model was introduced by Carr and Wun (2003). It is a Normal inverse Gaussian model that is stochastically time shifted by an integrated Cox-Ingersoll-Ross stochastic volatility process. The stock price process can be described

$$S_t = S_0 \exp(X_{I_t}), \tag{2.3}$$

where X_t is a NIG Lévy process, having parameters δ, α, β and $I_t = \int_0^t Y_s \mathrm{d}s$, Y_t has dynamics according to Equation (2.1) . Here the process $\{Y_t\}_{t \geqslant 0}$ is assumed to be independent of the process $\{X_t\}_{t \geqslant 0}$.

Given y_0 the characteristic function of Y_t can be found,

$$\begin{aligned} \phi_{CIR}(u,t,y_0) &= \mathbb{E}[\exp(\mathrm{i}uY_t)|y_0] \\ &= \frac{\exp(\kappa^2\eta t/\lambda^2)\exp(2y_0\mathrm{i}u/(\kappa+\gamma\coth(\gamma t/2)))}{(\cosh(\gamma t/2)+\kappa\sinh(\gamma t/2)/\gamma)^{2\kappa\eta/\lambda^2}}, \end{aligned}$$

where $\gamma = \sqrt{\kappa^2 - 2\lambda^2\mathrm{i}u}$. The risk-neutral price process S_t, see Equation (2.3), can be defined as,

$$S_t = S_0 \frac{e^{rt}}{\mathbb{E}[\exp(X_{Y_t})|y_0]} \exp(X_{Y_t}). \tag{2.4}$$

To make the derivation of the characteristic function for the log price of a NIG-CIR process easier we first express the expected value of the exponent of the NIG-CIR process:

$$\begin{aligned} \mathbb{E}[\exp(X_{Y_t})|y_0] &= \mathbb{E}[\mathbb{E}[\exp(X_{Y_t})|Y_t]|y_0] \\ &= \mathbb{E}[\phi_X(-\mathrm{i}, Y_t)|y_0] \\ &= \mathbb{E}[\exp(\kappa_{NIG}(1)Y_t)|y_0] \\ &= \phi_{CIR}(-\mathrm{i}\kappa_{NIG}(1), t, y_0) \end{aligned}$$

where the characteristic exponent $\kappa_{NIG}(u)$ is given by

$$\kappa_{NIG}(u) = ru - \delta(\sqrt{\alpha^2 - (\beta+u)^2} - \sqrt{\alpha^2 - \beta^2}).$$

The characteristic function for the log stock price, $\phi_{s_t}(u,t,y_0)$, is now easily derived by using the result above and writing out the whole expression for $\mathbb{E}[\exp(\mathrm{i}u\log(S_t))|S_0, y_0]$ according to the definition of S_t in Equation (2.4):

$$\phi_{s_t}(u,t,y_0) = \exp(\mathrm{i}u(rt + \log(S_0)) \frac{\phi_{CIR}(-\mathrm{i}\kappa_{NIG}(\mathrm{i}u), t, y_0)}{\phi_{CIR}(-\mathrm{i}\kappa_{NIG}(1), t, y_0)^{\mathrm{i}u}}.$$

2.2 Theoretical option price calculation using Inverse Fourier transformation

2.2.1 Carr-Madan Formula

Since the probability density of a Lévy processes is typically not known in closed form, there are no explicit formula for option prices in stochastic volatility with jumps models based on the Lévy processes. However, the characteristic function can be derived for some models which leads to the development of Fourier-based option pricing method. The Fourier transform based method have been frequently used in financial applications. For those who are interested in this method, Carr and Madan (1999) is a good reference to start with. A long list of references to articles using Fourier transform based methods can be found in e.g. Carr and Wun (2003).

Suppose t and T are current time and time of maturity, S_T and K are stock price at maturity date and strike price respectively. Set $s = \ln S_T$ and $k = \ln K$, then the payoff of the European call option can be written as

$$(S - K)^+ = (e^s - e^k)^+ \doteq f(k).$$

Let $C_T(k, t)$ denote the price of the European call option, then:

$$\begin{aligned} C_T(k,t) &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[f(k)|\mathcal{F}_t] \\ &= \int_k^\infty e^{-r(T-t)}(e^s - e^k)q_T(s)\mathrm{d}s, \end{aligned}$$

where q_T is the risk neutral density of $s = \ln S_T$.

To calculate the price of the call option under a certain price process we use Fourier Transforms suggested by Carr and Madan (1999). For this approach the characteristic function of the log-prices has to be given analytically which is the case for all of the most common model.

However, as the Fourier transform work for square-integrable functions only, we make the following modification, see Carr and Madan (1999) for details:

$$c_T(k) = e^{\alpha k}C_T, \quad \alpha > 0. \tag{2.5}$$

The Fourier transform of c_T is defined as:

$$\psi_T^C(\upsilon) = \int_{-\infty}^\infty e^{\mathbf{i}\upsilon k}c_T(k)\mathrm{d}k.$$

By inserting Equation (2.5) we get that:

$$
\begin{aligned}
\psi_T^C(\upsilon) &= e^{-r(T-t)} \int_{-\infty}^{\infty} e^{\mathbf{i}\upsilon k} \int_k^{\infty} e^{\alpha k}(e^s - e^k) q_T(s) \mathrm{d}s \mathrm{d}k \\
&= e^{-r(T-t)} \int_{-\infty}^{\infty} q_T(s) \int_{-\infty}^{s} e^{\mathbf{i}\upsilon k}(e^{\alpha k + s} - e^{(\alpha+1)k}) \mathrm{d}k \mathrm{d}s \\
&= e^{-r(T-t)} \int_{-\infty}^{\infty} q_T(s) \left(\frac{e^{(\alpha+1+\mathbf{i}\upsilon)s}}{\alpha + \mathbf{i}\upsilon} - \frac{e^{(\alpha+1+\mathbf{i}\upsilon)s}}{\alpha + \mathbf{i}\upsilon + 1} \right) \mathrm{d}s \\
&= \frac{e^{-r(T-t)} \phi_s(\upsilon - (\alpha+1)\mathbf{i})}{\alpha^2 + \alpha - \upsilon^2 + \mathbf{i}(2\alpha+1)\upsilon},
\end{aligned}
$$

where ϕ_s is the Fourier transform of q_T under a given asset model, i.e. $\phi_s = \mathbb{E}\left[e^{\mathbf{i}\upsilon \log S_T} | \mathcal{F}_t\right]$.

Hence the price of a European call option is given by the Fourier transform:

$$
C_T(k,t) = \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} e^{-\mathbf{i}\upsilon k} \psi_T^C(\upsilon) \mathrm{d}\upsilon.
$$

To simplify the expression we make a variable substittion by denoting $z = \mathbf{i}\upsilon + \alpha$. Furthermore, we use the notation where $X_T = \log(\frac{S_T}{S_t})$. As the characteristic function of X_T is written:

$$
\phi_{X_T|\mathcal{F}_t}(\upsilon, T) = \mathbb{E}\left[e^{\mathbf{i}\upsilon X_T} | \mathcal{F}_t\right] = \phi_s(\upsilon, T) e^{-\mathbf{i}\upsilon \log(S_t)},
$$

the price of a European call option can thus be written as:

$$
\begin{aligned}
C_T(k) &= \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} e^{-\mathbf{i}\upsilon k} \frac{e^{-r(T-t)} \phi_s(\upsilon - (\alpha+1)\mathbf{i}, T)}{\alpha^2 + \alpha - \upsilon^2 + \mathbf{i}(2\alpha+1)\upsilon} \mathrm{d}\upsilon, \\
&= \frac{e^{-r(T-t)}}{2\pi} \int_{\alpha - \mathbf{i}\infty}^{\alpha + \mathbf{i}\infty} \frac{e^{-kz} e^{(z+1)\log(S_t)} \phi_X(-\mathbf{i}(z+1), T)(-\mathbf{i})}{z(z+1)} \mathrm{d}z \\
&= \frac{e^{-r(T-t)}}{2\pi\mathbf{i}} \int_{\alpha - \mathbf{i}\infty}^{\alpha + \mathbf{i}\infty} \frac{e^{-kz} e^{(z+1)\log(S_t)} \phi_{X_T|\mathcal{F}_t}(-\mathbf{i}(z+1), T)}{z(z+1)} \mathrm{d}z \\
&= \frac{e^{-r(T-t)}}{2\pi\mathbf{i}} \int_{\alpha - \mathbf{i}\infty}^{\alpha + \mathbf{i}\infty} \frac{e^{-kz} e^{(z+1)\log(S_t)}}{z(z+1)} \mathbb{E}\left[e^{(z+1)X_T} | \mathcal{F}_t\right] \mathrm{d}z
\end{aligned}
$$

Define a function specific for a European call option with log strike price k:

$$
\Upsilon_C(z) = \frac{e^{-kz}}{2\pi\mathbf{i} z(z+1)}.
$$

The simplified expression of the call price yields,

$$C_T(k) = e^{-r(T-t)} \int_{\alpha-\mathbf{i}\infty}^{\alpha+\mathbf{i}\infty} \Upsilon_C(z) e^{(z+1)\log(S_t)} \phi_{X_T|\mathcal{F}_t}(-\mathbf{i}(z+1), T) \mathrm{d}z$$

For this to work the real part of z has to be positive and satisfy the condition $E^{\mathbb{Q}}[S_T^{1+Re(z)}] < \infty$, i.e. $Re(z) \in A_{X_T}^+ = \{x > 0 : E^{\mathbb{Q}}[S_T^{1+Re(z)}] < \infty\}$. According to the rule to choose $Re(z)$ in Lindström et al. (2008) Lindström et al. (2008), we can use the golden-section search method to find a proper value of $Re(z)$,

$$Re(z)_{\min} = \arg\min_{Re(z) \in A_{X_T}^+} g(Re(z)),$$

where

$$g(z) = \frac{e^{-zk+(z+1)s_t}}{z(z+1)} M_{X_T|\mathcal{F}_t}(Z+1),$$

and

$$M_{X_T|\mathcal{F}_t}(Z+1) = \phi_{X_T|\mathcal{F}_t}(-\mathbf{i}(z+1), T).$$

The Gauss-Laguerre quadrature formula is used to approximate an exponentially weighted integral from zero to infinity as

$$\int_0^\infty e^{-x} f(x) \mathrm{d}x \approx \sum_{j=1}^{n} \omega_j^{(n)} f(x_j^{(n)}).$$

More details can be found in Lindström et al. (2008).

2.2.2 Computation of the Delta

Most of the characteristic functions $\phi_T(\omega)$ that arise from stochastic volatility models can be rewritten as

$$\phi_T(\omega) = e^{\mathbf{i}\omega s_t} \times g(\omega),$$

where $g(\omega)$ is a function, that contains the remaining part of the characteristic function but not any s_t-terms. In this case, the derivative with respect to the initial spot price S_t is simply given by

$$\frac{\partial \phi_T(\omega)}{\partial S_t} = \frac{\partial \phi_T(\omega)}{\partial s_t} \frac{\partial s_t}{\partial S_t} = \frac{\mathrm{i}\omega}{S_t} \phi_T(\omega),$$

Thus, the derivative of the Carr-Madan formula for the call price $C_T(k)$ can be computed as

$$\begin{aligned}\frac{\partial C_T(k)}{\partial S_t} &= \frac{e^{-\alpha k}}{2\pi S_t}\int_{-\infty}^{\infty} e^{-\mathbf{i}\upsilon k}\frac{(\mathbf{i}\upsilon+\alpha+1)e^{-r(T-t)}\phi_s(\upsilon-(\alpha+1)\mathbf{i},T)}{\alpha^2+\alpha-\upsilon^2+\mathbf{i}(2\alpha+1)\upsilon}\mathrm{d}\upsilon,\\ &= e^{-r(T-t)}\int_{\alpha+1+i\mathbb{R}} \Upsilon_C(z)(z+1)e^{z\log S_t}M_X(z+1)\mathrm{d}z.\end{aligned}$$

2.3 Dynamic hedging strategies

A basic problem in Mathematical Finance is how the issue of an option can hedge the resulting exposure by trading in the underlying. Throughout this section we will denote the market price of the hedged contract $H(t, S_t)$. The price may contain other variables too, however as to simplify notation and increase readability, these variables have been exempted in the expression. First, we will start by assuming that we can use only the stock underlying the hedged option, and the risk-free asset as instruments in our hedge portfolio. Then we will add a call option as an additional instrument to the hedge portfolio. In this paper, the discrete time dynamic hedging strategies are considered.

2.3.1 Hedging using the underlying and the risk-free asset

The hedge portfolio in this case is properly defined,

$$P^h(t, S_t) = h_S S(t) + h_B B(t).$$

The superscript of $P(t, S_t)$, $h = (h_S, h_B)$, denote that $P(t, S_t)$ is determined by the vector of weights in each instrument. A hedge strategy is to choose the weights (h_S, h_B) is such a way that the portfolio $P^h(t, S_t)$ follows the derivative $H(t, S_t)$ as close as possible during a certain time interval.

Delta hedge A simple approach to hedge a financial contract $H(t, S_t)$ would be to set up a self-financing portfolio h such that:

$$\begin{cases}\dfrac{\partial P^h(t,S_t)}{\partial S} = \dfrac{\partial H(t,S_t)}{\partial S}\\ P^h(t,S_t) = H(t,S_t).\end{cases}$$

Thus, the hedging weight, $h = (h_S, h_B)$ is given by,

$$\begin{aligned}h_S^* &= \frac{\partial H(t,S_t)}{\partial S}\\ h_B^* &= H(t,S_t) - h_S^* S_t.\end{aligned}$$

We may approximate the *delta* $\Delta = h_S^*$ by finite difference ratios,

$$\frac{\partial H(t, S_t)}{\partial S} = \lim_{\Delta S \to 0} \frac{H(t, S_t + \Delta S) - H(t, S_t - \Delta S)}{2\Delta S},$$

or by differentiating the price expression with respect to S_t, in the call option case, such that:

$$\frac{\partial H(t, S_t)}{\partial S} = e^{-r(T-t)} \int_{\alpha+1+i\mathbb{R}} \Upsilon_C(z)(z+1) e^{z \log S_t} M_X(z+1) \mathrm{d}z.$$

Quadratic hedging Quadratic hedging approaches have been studied very intensively in recent decades. It is a popular technique that produces hedges that minimize the quadratic hedging error in mean square sense. One of the reasons that quadratic hedging is popular among scholars and practitioners is that it can be formulated in a mathematically elegant way and solved by relatively easy approaches.

An obvious drawback of quadratic hedging is that losses and gains are treated in the same way. On the other hand, this might be an advantage, in case you do not know whether you deal with a buyer or a seller. Another advantage is that quadratic strategies related to different options can simply be added up as is also the case for delta-hedging strategies. In other words, quadratic hedging is a sort of linear hedging strategy.

In the Black-Scholes model (Black and Scholes, 1973), the market is assumed to be complete so that a perfect hedge can be attained by delta hedging. However, generally in an incomplete market, a perfect hedge is not attainable. Foellmer and Sondermann (1986) formulate the risk-minimization strategy, which is a breakthrough in quadratic hedging in incomplete market, see also Föllmer and Schweizer (1989) and Schweizer (1991). The strategy adopts a mean-self-financing strategy rather than self-financing strategy to minimize the expected future costs.

Define V_t as the value of the hedge portfolio at a discrete time point t after rebalancing. The value of the same portfolio at the next rebalance point is denoted $V_{t+\Delta t}$. S_t is the value of the spot price at time point t while B_t is the risk free asset. $S_{t+\Delta t}$ is the spot value at time point $(t + \Delta t)$, i.e. a time of Δt has passed since time point t. The value of portfolio at time t and Δt have the expressions as follows:

$$\begin{aligned} V_t &= h_{S_t} S_t + h_{B_t} B_t, \\ V_{t+\Delta t} &= h_{S_t} S_{t+\Delta t} + h_{B_t} B_t e^{r\Delta t} \\ &= V_t e^{r\Delta t} + h_{S_t}(S_{t+\Delta t} - e^{r\Delta t} S_t). \end{aligned}$$

By definition $V_t = H_t$, which leads to a definition of the quadratic

hedging error, R, over a given time period Δt as:

$$R = (H_{t+\Delta t} - V_{t+\Delta t})^2 = \left((H_{t+\Delta t} - e^{r\Delta t}H_t) - h_{S_t}(S_{t+\Delta t} - e^{r\Delta t}S_t)\right)^2. \tag{2.6}$$

As mentioned above, we wish to minimize R with respect to h_{S_t} under either physical measure $\mathbb{P}$ or risk neutral measure $\mathbb{Q}$ such that:

$$\begin{aligned} h_S^* &= \arg\min_{h_S} \mathbb{E}[R|\mathcal{F}_t] \\ &= \arg\min_{h_S} \mathbb{E}[\left((H_{t+\Delta t} - e^{r\Delta t}H_t) - h_{S_t}(S_{t+\Delta t} - e^{r\Delta t}S_t)\right)^2 |\mathcal{F}_t] \\ &= \arg\min_{h_S} \mathbb{E}[(\Delta H)^2 - 2h_s\Delta H\Delta S + h_s^2(\Delta S)^2|\mathcal{F}_t]. \end{aligned}$$

where $\Delta H = H_{t+\Delta t} - e^{r\Delta t}H_t$ and $\Delta S = S_{t+\Delta t} - e^{r\Delta t}S_t$. Expanding the square of the expression and taking the expected value yields the convex minimization with the straight forward solution:

$$h_S^* = \frac{\mathbb{E}[\Delta H\Delta S|\mathcal{F}_t]}{\mathbb{E}[(\Delta S)^2|\mathcal{F}_t]}. \tag{2.7}$$

By consideration Equation 2.7 under risk neutral measure $\mathbb{Q}$, we get $\mathbb{E}^{\mathbb{Q}}[\Delta S|\mathcal{F}_t] = 0$ and $\mathbb{E}^{\mathbb{Q}}[\Delta H\Delta S|\mathcal{F}_t] = \mathrm{Cov}(\Delta H, \Delta S)$, thus the optimal portfolio under risk neutral measure $\mathbb{Q}$ can be written as,

$$h_S^{\mathbb{Q}} = \frac{\mathrm{Cov}^{\mathbb{Q}}(S_{t+\Delta t}, H_{t+\Delta t}|\mathcal{F}_t)}{\mathrm{Var}^{\mathbb{Q}}(S_{t+\Delta t}|\mathcal{F}_t)}.$$

The optimal portfolio is readily found if the expectation is approximated by a Monte Carlo sample, as the optimization problem is transformed into a standard least squares regression. Sampling N particles from $\pi^{\mathbb{P}}(S_{t+1}|S_t)$ gives

$$\begin{aligned} h_S^{\mathbb{P}} &= \arg\min_{h_S} \frac{1}{N}\sum_{n=1}^{N}\left[\left((H_{t+\Delta t}^{(n)} - e^{r\Delta t}H_t) - h_{S_t}(S_{t+\Delta t}^{(n)} - e^{r\Delta t}S_t)\right)^2 |\mathcal{F}_t^{\mathbb{P}}\right] \\ &= \frac{\frac{1}{N}\sum_{n=1}^{N}\left((H_{t+\Delta t}^{(n)} - e^{r\Delta t}H_t)(S_{t+\Delta t}^{(n)} - e^{r\Delta t}S_t)\right)}{\frac{1}{N}\sum_{n=1}^{N}\left(S_{t+\Delta t}^{(n)} - e^{r\Delta t}S_t\right)^2}. \end{aligned}$$

2.3.2 Adding another hedge instrument

Now, we augment the conditions for hedging a derivative by adding a call option as an extra hedge instrument. Thus, we define our hedge portfolio the following way:

$$P^h(t, S_t) = h_S S(t) + h_C C(t, S(t)) + h_B B(t).$$

The instrumental call option is denoted $C(t, S(t))$. The augmentation should lead to better hedge schemes than those we have presented before. It is important that the time to maturity for the instrumental option is longer than that of the hedged option. Similar to Equation (2.6) we define the quadratic hedging error R, as a random variable standing at time point t, looking forward a short time interval Δt.

$$R = ((H_{t+\Delta t} - e^{r\Delta t} H_t) - h_{S_t}(S_{t+\Delta t} - e^{r\Delta t} S_t) - h_{C_t}(C_{t+\Delta t} - e^{r\Delta t} C_t))^2.$$

By expanding the square and taking the expected value of the expression we are left with a convex two-dimension minimization problem. The asset weights h_S and h_C that correspond to the optimal hedge portfolio physical measure $\mathbb{P}$ is:

$$h_S^{\mathbb{P}} = \frac{\mathbb{E}[\Delta C^2]\mathbb{E}[\Delta S \Delta H] - \mathbb{E}[\Delta S \Delta C]\mathbb{E}[\Delta C \Delta H]}{\mathbb{E}[\Delta S^2]\mathbb{E}[\Delta C^2] - \mathbb{E}^2[\Delta S \Delta C]},$$
$$h_C^{\mathbb{P}} = \frac{\mathbb{E}[\Delta S^2]\mathbb{E}[\Delta C \Delta H] - \mathbb{E}[\Delta S \Delta C]\mathbb{E}[\Delta S \Delta H]}{\mathbb{E}[\Delta S^2]\mathbb{E}[\Delta C^2] - \mathbb{E}^2[\Delta S \Delta C]},$$

where $\Delta H = H_{t+\Delta t} - e^{r\Delta t} H_t$, $\Delta S = S_{t+\Delta t} - e^{r\Delta t} S_t$ and $\Delta C = C_{t+\Delta t} - e^{r\Delta t} C_t$. Under risk neutral measure $\mathbb{Q}$, the expected value of ΔH, ΔS and ΔC are all zero, thus the optimal hedge portfolio are:

$$h_S^{\mathbb{Q}} = \frac{\mathrm{Var}(C)\mathrm{Cov}(S, H) - \mathrm{Cov}(S, C)\mathrm{Cov}(C, H)}{\mathrm{Var}(S)\mathrm{Var}(C) - \mathrm{Cov}^2(S, C)},$$
$$h_C^{\mathbb{Q}} = \frac{\mathrm{Var}(S)\mathrm{Cov}(C, H) - \mathrm{Cov}(S, C)\mathrm{Cov}(S, H)}{\mathrm{Var}(S)\mathrm{Var}(C) - \mathrm{Cov}^2(S, C)}.$$

All means, variances and co-variances are conditioned on the filtration $\mathcal{F}_t$, and subscripted $t + \Delta t$.

3 Monte Carlo Simulation

In the current section we describe in some detail how the particular processes presented in model calibration and can be implemented in practice in a Monte Carlo simulation pricing framework. For this we discuss the numerical implementation of the stochastic volatility jump diffusion model and then introduce three building block processes which drive them. This will be followed by an explanation of how one assembles a time changed Lévy process.

3.1 Simulation of the Heston Model

For numerical purposes, it is preferred to simulate the change in the logarithm of the asset price. We define $X_t = \ln S_t$ and apply Itô's lemma under risk neutral measure to get

$$\mathrm{d}X_t = (r - \frac{1}{2}V_t)\mathrm{d}t + \sqrt{V_t}\mathrm{d}W_t^{\mathbb{Q}}. \tag{3.1}$$

A discrete-time approximation of Equations 3.1 and 2.1 using the Euler-Maruyama scheme is given by

$$X_{t+\Delta t} = X_t + (r - \frac{1}{2}V_t)\Delta t + \sqrt{V_t}\sqrt{\Delta t}Z_t^1,$$

$$V_{t+\Delta t} = V_t + \kappa(\xi - V_t)\mathrm{d}t + \sigma\sqrt{V_t}\sqrt{\Delta t}Z_t^2. \tag{3.2}$$

where Z_t^1 and Z_t^1 are two standard normal random variables with correlation ρ. We use the Euler-Maruyama scheme in simulation through this paper.

3.2 Simulation of the Bates model

The independence of the Poisson process, the jump size and the Brownian motion driving the underlying strongly facilitates the discretisation of the Bates model. As a advantage of the Heston model with jump, it is possible to split up the computation of X_t into a geometric Brownian motion part, a drift adjustment for the jump component and the value of the compound Poisson process. Thus the simulation steps of Bates model can be summarized as:

1. sample the number of jumps n_J from Poisson distribution with intensity $\lambda\Delta t$,

2. sample the jumps size J from Normal distribution with mean μ and variance δ^2,

3. calculate the drift part $-\lambda(\exp(\frac{\delta^2}{2}+\mu)-1)$.

4. a discrete-time approximation of X_t using the Euler-Maruyama scheme is given by

$$X_{t+\Delta t} = X_t + (r - \frac{1}{2}V_t + \text{drift})\Delta t + \sqrt{V_t}\sqrt{\Delta t}Z_t^1 + \sum_{j=0}^{n_J} J_j.$$

3.3 Simulation of the time-changed Lévy model

3.3.1 NIG Lévy Process

To simulate a NIG process, we first describe how to simulate NIG(α, β, δ) random numbers. NIG random numbers can be obtained by mixing Inverse Gaussian (IG) random numbers and standard Normal numbers in the following manner. An IG(a, b) random variable X has a characteristic function given by:

$$\mathbb{E}[\exp(\mathrm{i}uX)] = \exp(-a(\sqrt{-2u\mathrm{i}+b^2}-b)).$$

First simulate IG(1,δ,$\sqrt{\alpha^2-\beta^2}$) random numbers i_k, for example using the Inverse Gaussian generator of Michael, Schucany and Haas (Devroye and Devroye, 1986). Then sample a sequence of standard Normal random variables u_k. NIG random numbers n_k are then obtained via:

$$n_k = \delta^2\beta i_k + \delta\sqrt{i_k}u_k.$$

Finally the sample paths of a NIG(α, β, δ) process $X = \{X_t, t \geqslant 0\}$ in the time points $t_n = n\Delta t, n = 0, 1, 2, ...$ can be generated by using the independent NIG(α, β, $\delta\Delta t$) random numbers n_k as follows:

$$X_0 = 0, \qquad X_{t_k} = X_{t_{k-1}} + n_k, \quad k \geqslant 1.$$

3.3.2 CIR Stochastic Clock

The simulation of a CIR process $y = \{y_t, t \geqslant 0\}$ is straightforward. Basically, we discretize the SDE:

$$\mathrm{d}y_t = \kappa(\eta - y_t)\mathrm{d}t + \lambda y_t^{1/2}\mathrm{d}W_t, \quad y_0 \geqslant 0,$$

where W_t is a standard Brownian motion. Using a first-order accurate explicit differencing scheme in the time points $t_n = n\Delta t, n = 0, 1, 2, ...$, the sample path of the CIR process $y = \{y_t, t \geqslant 0\}$ is then given by Equation 3.2.

3.3.3 Path Generation for Time Changed Lévy Process

The explanation of the building block processes as stated above allow us to assemble all the parts of the time-changed Lévy process simulation puzzle. For this one can proceed through the following five steps (Schoutens, 2003).

1. simulate the rate of time change process $y = \{y_t, 0 \leqslant t \leqslant T\}$,
2. calculate the time change $Y = \{Y_t = \int_0^t y_s \mathrm{d}s, 0 \leqslant t \leqslant T\}$,
3. simulate the Lévy process $X = \{X_t, 0 \leqslant t \leqslant Y_T\}$,
4. calculate the time changed Lévy process X_{Y_t}, $0 \leqslant t \leqslant T$,
5. calculate the stock price process using Equation 2.4. The correcting factor is calculated as:

$$\frac{e^{rt}}{\mathbb{E}[\exp(X_{Y_t})|y_0]} = \frac{e^{rt}}{\phi_{CIR}(-\mathrm{i}\kappa_X(-\mathrm{i}), t, 1)}.$$

4 Model calibration

A stable calibration of option pricing models is of paramount importance for investment banks because the price of options are determined by estimated models. This section will review standard calibration methods, discuss and compare a few new technologies leading to better use of data.

4.1 Least squares methods

The dominating calibration method is the weighted least square method (Cont and Tankov, 2004; Schoutens et al., 2004; Hull, 2009), i.e. taking the parameter vector of the model that minimizes the weighted sum of the squared difference between the observed mid-price and model price

$$\begin{aligned} L_t^{WLS}(\theta) &= \sum_{s=1}^{t}\sum_{i=1}^{N_s} \lambda_{s,i}\left(c_s^{\star}(K_i,\tau_i) - c_s^{Model}(K_i,\tau_i;\theta)\right)^2, \\ \hat{\theta}_t &= \arg\min_{\theta\in\Theta} L_t^{WLS}(\theta). \end{aligned}$$

It is statistically optimal 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 $\lambda_{s,i}$ 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

$$\lambda_{s,i} \propto \frac{1}{\left(c_s^{Ask}(K_i,\tau_i) - c_s^{Bid}(K_i,\tau_i)\right)^2}. \tag{4.1}$$

Since this calibration is known to be numerically difficult to determine the parameters minimizing the loss function and has problem of easily over fitting data with a small set of observations, a couple of alternative approaches have emerged over the last few years.

4.2 Non-linear filtering

In this paper, EnKF is used to compute the distribution of the latent states given the observations according to the above mentioned analysis. The updating of the distribution of the latent states is based on the theory on linear projections in $\mathbb{L}^2$(best linear estimate), similar to what is done in the EKF and IEKF (Lindström et al., 2008).

Consider a discrete-time nonlinear system with dynamics

$$x_t = f(x_{t-1}) + w_t,$$

and measurements

$$y_t = h(x_t, u_t) + \epsilon_k,$$

where x_t, $w_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^m$, y_t, and $\epsilon_t \in \mathbb{R}^p$. We assume that w_t and v_t are stationary zero-mean white noise processes with covariance matrices Q_t and R_t, respectively. Furthermore, we assume that x_0, w_t and v_t are uncorrelated. The model can also depend on a known (control) input u_t, which represent the set of $\mathcal{F}_t$ measurable variables (e.g. $S_t, K_t, \tau_t, \cdots$).

The algorithm of EnKF is summarized as following,

Ensemble Kalman filter

Initialisation: Let $p(x_0) \sim N(m_0, P_0)$ and draw N samples $x_0^{(n)}$.

Propagate: Propagate the latent state as

$$\begin{aligned} x_{t|t-1}^{(n)} &\sim p(X_t | x_{t-1|t-1}^{(n)}) \\ y_{t|t-1}^{(n)} &= h(x_{t|t-1}^{(n)}) \end{aligned}$$

and compute

$$\begin{aligned} \overline{x}_{t|t-1} &= \frac{1}{N}\sum x_{t|t-1}^{(n)},\ \overline{y}_{t|t-1} = \frac{1}{N}\sum y_{t|t-1}^{(n)} \\ E_x &= [x_{t|t-1}^{(1)} - \overline{x}_{t|t-1} \ \cdots \ x_{t|t-1}^{(N)} - \overline{x}_{t|t-1}] \\ E_y &= [y_{t|t-1}^{(1)} - \overline{y}_{t|t-1} \ \cdots \ y_{t|t-1}^{(N)} - \overline{y}_{t|t-1}] \\ P_{xy} &= \frac{1}{N-1} E_x E_y^T,\ P_{yy} = \frac{1}{N-1} E_y E_y^T \end{aligned}$$

where R is the covariance of ε.

Updating: The updated particle representation of the filter density is given by

$$\begin{aligned} K_t &= P_{xy}(P_{yy} + R)^{-1} \\ x_{t|t}^{(n)} &= x_{t|t-1}^{(n)} + K_t(y_t + \epsilon^{(n)} - y_{t|t-1}^{(n)}) \end{aligned}$$

where $\epsilon^{(n)} \overset{\mathrm{d}}{=} \varepsilon$.
and repeat the propagation and updating steps until $t = T$.

The ensemble Kalman filter moves that particles, using linear trans-

formations rather than resampling particles as is done in particle filters, thus avoiding sample degeneration.

4.3 An empirical study of model calibration and quadratic hedging

4.3.1 Data description

We have used daily data on S&P 500 index options, from November 5th, 2001 to May 5th, 2003. The original data set consisted of 38225 quotes (date, strike, ask price, bid price, index level, time to maturity and risk free interest rate) and there are 376 trading days for the S&P 500 index. In order to reduce the problems with infrequent trading and other noises in the market, we have applied several data selection criteria,

1. The options with a maturity of fewer than 6 days and longer than 1 year have been excluded from the investigation. The options with a maturity less than one week have relatively small time premiums, and contain very little information about the volatility. Options with maturity longer than one year are not actively traded.

2. Options having zero or negative bid-ask spread and bid-ask spread larger than 5 have been excluded. Also options with prices less than \$0.01 have been eliminated.

3. The least liquid option when call and put options having identical strike and time to maturity is available, have been eliminated.

After this processing, there are 238 trading days, 27978 valid unique quotes left. We transform all quoted put options into call options using put-call parity. The statistics of the data set can be found in Table 1.

	Moneyness S/K	Days to Expiration < 60	$60 - 180$	$\geqslant 180$	Subtotal
		\$5.11	\$16.67	\$40.34	
	< 0.94	(0.35)	(0.50)	(0.61)	
		{1712}	{1656}	{101}	{3469}
OTM		\$12.25	\$26.47	\$49.68	
	0.94-0.97	(0.46)	(0.52)	(0.62)	
		{2061}	{2416}	{575}	{5052}
		\$23.71	\$39.96	\$62.89	
	0.97-1.00	(0.50)	(0.53)	(0.62)	
		{2005}	{2356}	{755}	{5116}
ATM		\$40.10	\$56.17	\$78.80	
	1.00-1.03	(0.50)	(0.53)	(0.62)	
		{1815}	{2168}	{676}	{4659}
		\$59.47	\$74.08	\$90.30	
	1.03-1.06	(0.49)	(0.52)	(0.63)	
		{1725}	{1992}	{314}	{4031}
ITM		\$93.95	\$99.49	\$112.35	
	$\geqslant 1.06$	(0.42)	(0.51)	(0.59)	
		{3058}	{2499}	{94}	{5651}
Subtotal		{12376}	{13087}	{2515}	{27978}

Table 1: Summary statistics for data set. The reported numbers are respectively the average quoted bid-ask mid point price, the average half of bid-ask spread, which shown in parentheses, and the total number of observations (in braces), for each moneyness-maturity category. S denotes the spot S&P 500 index level and K is the strike price. OTM, ATM, and ITM denote out-of-the-money, at-the-money, and in-the-money,respectively.

4.3.2 Implementation

Lindström et al. (2008) suggested that parameters and the latent volatility can be estimated recursively by applying a sequential filter to

$$\begin{aligned} c_t^\star(K_i, \tau_i) &= c^{Model}(K_i, \tau_i; \widetilde{\theta}_t, V_t) + \epsilon_t, \\ \widetilde{\theta}_t &= \widetilde{\theta}_{t-1} + \eta_t^\theta, \\ V_t &\sim \pi^{\mathbb{P}}(V_t | V_{t-1}; \widetilde{\theta}_{t-1}), \end{aligned} \tag{4.2}$$

where the first equation is interpreted as the measurement equation $Y_{i,t} = c_t^\star(K_i, \tau_i)$ and the latter two equations $X_t = [\theta_t \ V_t]^T$ are interpreted as latent processes. We assume that $\eta_t^{\theta,V}$, and ϵ_t are stationary zero-mean white noise processes with covariance matrices Q_t and R_t, respectively. The measurement noise covariance has a standard deviation of one fourth

of the spread, cf. $\mathbf{R} = \text{cov}(\epsilon) = (c^{Ask} - c^{Bid})^2/16$.The motivation behind this choice is that we expect at least 95% of the true option prices to be within the Ask-Bid spreads. The process noise covariance matrix Q was determined as $Q^c = 5 \cdot 10^{-3}$ and $Q^J = 10^{-4}$ with relation to the continuous component and jump component respectively.

This is a dynamic Bayesian method, addressing the problem of time varying parameters as well. The latent state vector V_t is organized with a transformed version of the parameter vector $\widetilde{\theta}_t = \alpha(\theta_t)$, where the transformed parameter vector is modeled as a random walk. Neglecting to transform the parameters may cause the algorithm to break down (Lindström et al., 2008). 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 $[-1, 1]$.

There is no need to specify the initial distribution of the latent states when apply the filter algorithm to data. Although we have initialized the filter using the asymptotics given by the WLS estimator, i.e. taking $x_0 = \theta_1^{WLS}$ and $P_0 = \text{Cov}(\theta_1^{WLS})$. This is a reasonable prior and ensures that the filter has good starting values.

The latent volatility and underlying dynamics mentioned above for different models are presented under the risk-neutral measure. Thus the parameters involved are $\mathbb{Q}$-parameters. In order to hedge the contract in the real world, the $\mathbb{P}$-parameters, which are under the physical measure, are needed. There is a relationship between $\mathbb{P}$ and $\mathbb{Q}$ parameters through risk premiums λ_V,

$$\kappa^{\mathbb{Q}} = \kappa^{\mathbb{P}} + \lambda_V, \quad \xi^{\mathbb{Q}} = \frac{\kappa^{\mathbb{P}}\xi^{\mathbb{P}}}{\kappa^{\mathbb{P}} + \lambda_V}.$$

We first calibrate risk-neutral measure so that the $\mathbb{Q}$-parameters are obtained. Then we test different volatility risk premie to match historical return. After doing this, there is no significant estimation was obtained. One possible reason is that the risk premia has tiny influence on prediction since we use the daily calibrated parameters to predict the next day index level and volatility dynamics. Thus for computational convenience, we assume the parameters of the model under $\mathbb{P}$ and $\mathbb{Q}$ dynamics are the same through this paper (Lindström et. al, 2008).

The models used in Monte Carlo hedging are not required to be the same as in calibration. However, we assume that the models are identical for calibration and hedging. In the hedging part, the transaction costs are not considered since not only do they tend to vary over time but they also depend on the size of the transaction.

Applying a non-linear Kalman filter (Lindström et al., 2008) uses an

iterated Extended Kalman filter) to the model generates and estimates an estimate of the latent parameters and states conditional on the observations

$$p(X_t|Y_{i,s} \ \forall \ s \leqslant t, \ \forall i).$$

A limitation when using an iterated Extended Kalman filter is that all distributions are treated as if they are Gaussian, something that may be inaccurate for jump processes. This limitation can be circumvented by Monte Carlo simulation of the latent processes. The advantage of using simulation is that we can generate an arbitrary number of paths, making the approximation error arbitrarily small.

The updating of the latent processes can without loss of performance use a Kalman filter type method as the dimension of the observation vector (i.e. the number of options quoted) usually exceed the dimension of the latent states. This makes the linearization of the measurement equation uncontroversial.

4.3.3 Results

In this section, the estimated parameters of the Bates model and the NIGCIR model from different calibration technique are first compared to test over-fitting. Then, the relationship between estimation efficiency and number of ensembles is illustrated by applying two hedging strategies on the estimations.

The calibrated parameters

The Bates model and NIGCIR model have been estimated on market data from options on the SP 500 index. 80% of the data from the above data set are chosen randomly and used for estimation and the remainder 20% are for validation. The parameter estimates of the Bates model using different calibration techniques, namely WLS, UKF and EnKF(with 400 ensembles), are presented in Figure 1, whereas the estimations of the NIGCIR model are shown in Figure 2.

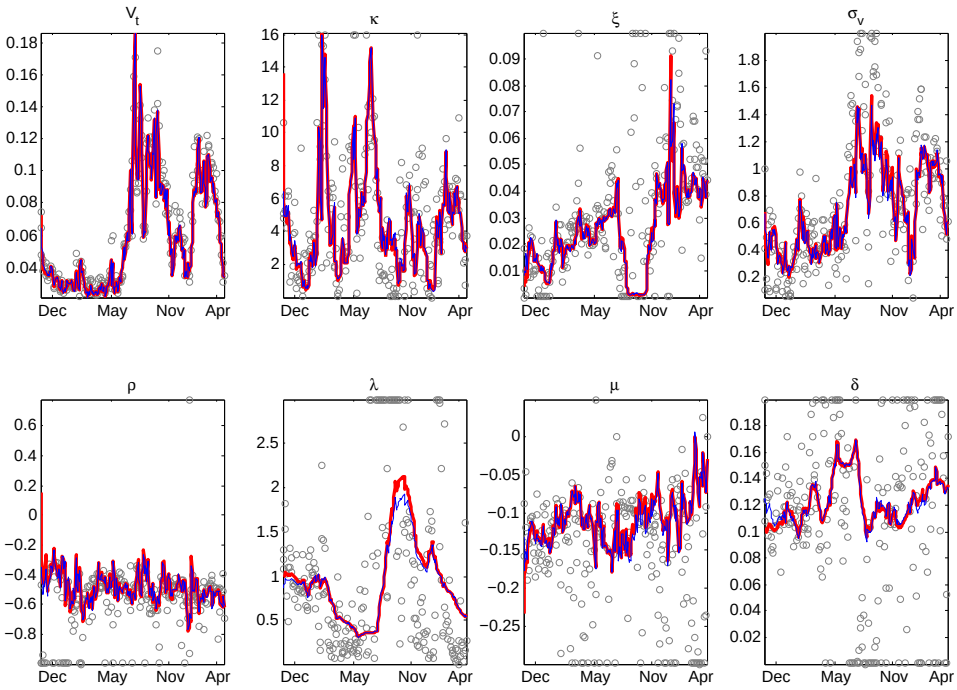

Figure 1: Bates on market data. WLS estimates (circles), UKF parameters (thin line), and EnKF estimates with 400 ensembles (thick line).

For the Bates model, the estimations from UKF and EnKF give similar patterns for all parameters, whereas the WLS estimates varies a great deal, except for the latent state process, the volatility V_t. There are some extreme estimates from WLS for the jump parameters, the jump intensity λ, the expected size μ and the standard deviation δ of the jumps. This shows that the parameters which measure jumps, is hard to track. But these slightly difference in jump parameters do not affect the model predicability. This is also seen in Table 2.

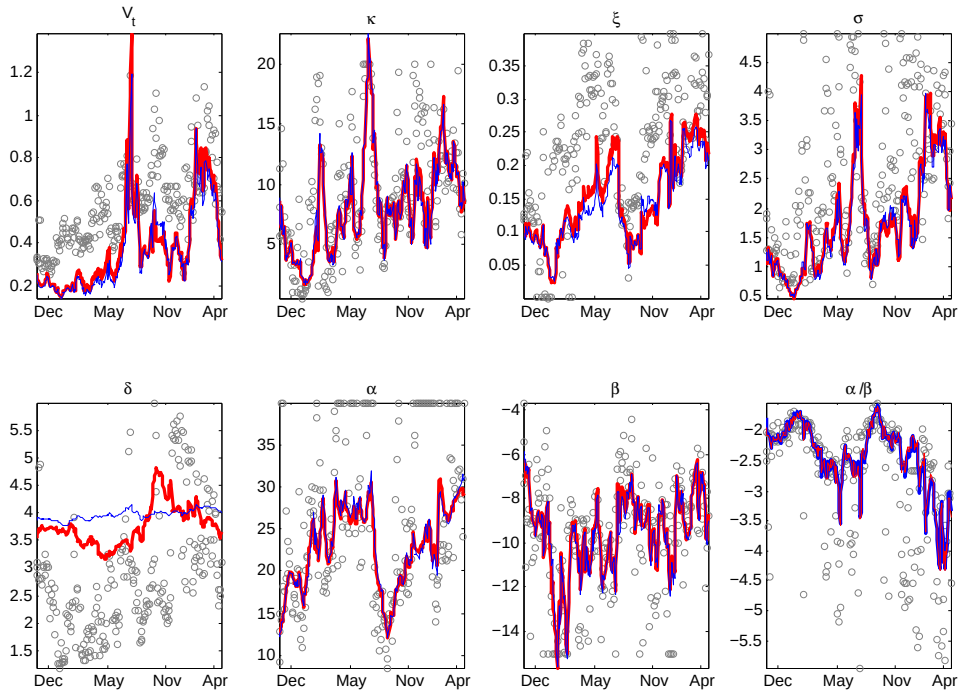

Figure 2: NIGCIR on market data. WLS estimates (circles), UKF parameters (thin line), and EnKF estimates with 400 ensembles (thick line).

For the NIGCIR model, the UKF estimates and EnKF estimates are closely to each other, except for the the scale parameter δ of the jumps. The WLS can't achieve stable and convergence parameter for the NIGCIR model for this data set as extreme values appear in the figure of parameter α, which controls the tail heaviness of the NIG distribution.

The accuracy of the model parameters directly affects the theoretical price forecast and the hedging strategy. In order to examine the models forecast performance, we have calculated the Root Mean Square Error (RMSE) for both in-sample and out-of-sample forecasts as following,

$$\mathrm{RMSE}(x) = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(\hat{x}_t - x_t)^2} = \sqrt{\frac{1}{N}\sum_{t=1}^{N} e_t^2},$$

where x_t is the mid-point of the bid and ask prices at time t, $\hat{x}_t$ is the predict price with the model calibrated parameters, i.e. $\hat{x}_t = \hat{x}_{t|t}$.

The performance of models has also been examined by the proportion of options price inside the bid-ask spread(IS):

$$\mathrm{IS}(x) = \frac{1}{N}\sum_{t=1}^{N} 1_{[\mathrm{bid,ask}]}(\hat{x}_t).$$

All of the model performance results of the RMSE and the IS from various methods are presented in table 2. As can be seen from the table, all methods perform well both on the in-sample data and the out-of-sample data. Both EnKF and UKF work well on this dataset. The WLS performs best on the Bates model, but for the more complex model, the NIGCIR model, the measurements shows the over-fitting of WLS. The measurement of EnKF with different numbers of ensembles do not change significantly while the ensemble size goes up. Thus 400 ensembles is sufficient enough for calibration for this data set.

Filter	Set	Bates		NIG-CIR	
		IS (%)	RMSE	IS (%)	RMSE
WLS	Est.	**89.66**	**0.2703**	86.18	0.3257
	Val.	**86.81**	**0.3048**	84.66	0.3497
UKF	Est.	87.09	0.3153	84.12	0.3525
	Val.	85.84	0.3356	82.84	0.3682
EnKF	Est.	87.95	0.3080	86.46	0.3269
(50 ensembles)	Val.	86.18	0.3236	84.77	0.3448
EnKF	Est.	88.09	0.3163	86.66	0.3250
(100 ensembles)	Val.	86.08	0.3322	84.72	0.3431
EnKF	Est.	88.57	0.2946	**86.89**	**0.3236**
(400 ensembles)	Val.	86.38	0.3191	**84.97**	**0.3417**

Table 2: Summary statistics for the calibration results of the Bates model and the NIGCIR model using different estimation methods. The reported numbers are the global fit proportional option prices inside the bid-ask spread(IS%) and error measurements for the in-sample (estimation set) and out-of-sample (validation set) of S&P 500 data set. Bold face indicates the best method for that measure.

Dynamic Hedging Strategy

The key point in a hedging strategy or any trading strategy with risk-neutral purpose is to obtain the weight of the hedging portfolio, and the weights differ from the ways how they measure the risk. We follow the hedging strategy as described in Section 2.3, namely Delta hedge and Monte Carlo quadratic hedging using underlying and risk-free asset under the risk neutral measure $\mathbb{Q}$. The trading step behind these two hedging strategies is one who buy/sell a group of options today, want to sell/buy them on the next observation day. The purpose of doing this is to find whether a relationship exists between the hedging efficiency and ensemble size.

	Par. Set	Bates	NIG-CIR
Δ-hedge	A	\$6.47(4.63)	\$7.01(5.54)
	B	\$6.48(4.63)	\$7.02(5.55)
	C	\$6.47(4.62)	\$7.02(5.55)
100 simulations	A	\$6.24(4.83)	\$6.24(4.90)
	B	\$6.21(4.98)	\$6.19(4.90)
	C	\$6.02(4.69)	\$6.03(4.76)
400 simulations	A	\$5.79(4.62)	\$6.06(4.76)
	B	\$5.62(4.57)	\$6.01(4.71)
	C	\$5.57(4.03)	\$5.95(4.74)
1000 simulations	A	\$5.58(4.13)	\$6.01(4.79)
	B	\$5.49(4.10)	\$5.94(4.76)
	C	\$5.33(4.11)	\$5.92(4.72)

Table 3: The hedging results for the Bates model and the NIG-CIR model of two different hedging strategies: Δ hedge and quadratic minimization. Parameter Set A indicate the EnKF estimates with 50 ensembles, while B and C is for EnKF estimates with 100 and 400 ensembles, respectively. The reported numbers are the values of average absolute hedging error over the whole period and the standard deviation(in braces).

The hedging results for Delta hedging and Monte Carlo quadratic hedging(with different number of simulations) are shown in Table 3. On one hand, one may expect the level of hedging error would decrease with the increase of the number of simulations. The measurements in the above table represent the same expectation since the errors for the same parameter set decline as the number of simulations goes up. However to find the ensemble size with which we could get the smallest hedging error during hedging, is not our purpose in this paper. On the other hand, the influence on hedging results with the changing of ensemble size doesn't show up. Note that the Monte Carlo quadratic hedging performs better than the standard Δ-hedging even with 100 Monte Carlo simulations for each parameter set.

5 Simultaneous calibration and hedging

We will modify the model in Section 6 by augmenting the latent processes with the underlying asset S_t and simultaneously including it in the measurement equation. The reason for doing so is that samples from the underlying asset is required when doing the hedging.

Introducing the augmented latent processes $X_t = [S_t,\ V_t,\ \theta_t]$ suggests that model 4.2 can be rewritten as

$$\begin{bmatrix} S_t^\star \\ c_t^\star(K_i, \tau_i) \end{bmatrix} = \begin{bmatrix} X_t(1) \\ c^{Model}(K_i, \tau_i; X_t) \end{bmatrix} + \begin{bmatrix} \epsilon_t^S \\ \epsilon_t^c \end{bmatrix}$$
$$X_t \sim \pi^{\mathbb{P}}(X_t | X_{t-1}),$$

where the first two equations are the measurement equations $Y_t^\star = [S_t^\star\ c_t^\star(K_i, \tau_i)]^T$ and the last equation is the augmented latent processes. The variance (derived from the ask-bid spread) of the measurement noise for underlying asset is typically much smaller than the variance of any of the options, ensuring that the filter estimate of the underlying asset is within the ask-bid spread for the asset.

5.1 Algorithm

Noting that the filter computed $X_{t|t-1} = [S_{t|t-1}, V_{t|t-1}, \theta_{t|t-1}]^T$ and $Y_{t|t-1} = [X_{t|t-1}(1); c^{Model}(K_i, \tau_i; X_{t|t-1})]$ as part of the propagation step makes quadratic hedging inexpensive to compute. The regression uses $\zeta_i = c^{Model}(K_i, \tau_i; X_{t|t-1})$ and $\Xi = [B_{t|t-1}, S_{t|t-1}]$, all of them already computed. Most importantly, the expensive computation of option values is already done!

We summarize the algorithm, integrating the filter and hedge regression below.

Simulanteous calibration and hedging

Initialisation: Let $p(x_0) \sim N(m_0, P_0)$ and draw N samples $x_0^{(n)}$.

Propagate: Propagate the latent state as

$$\begin{aligned} x_{t|t-1}^{(n)} &\sim \pi^{\mathbb{P}}(X_t | x_{t-1|t-1}^{(n)}) \\ y_{t|t-1}^{(n)} &= c^{Model}(x_{t|t-1}^{(n)}) \end{aligned}$$

and compute

$$\overline{x}_{t|t-1} = \frac{1}{N}\sum x_{t|t-1}^{(n)}, \quad \overline{y}_{t|t-1} = \frac{1}{N}\sum y_{t|t-1}^{(n)},$$

$$E_x = [x_{t|t-1}^{(1)} - \overline{x}_{t|t-1} \ \cdots \ x_{t|t-1}^{(N)} - \overline{x}_{t|t-1}],$$

$$E_y = [y_{t|t-1}^{(1)} - \overline{y}_{t|t-1} \ \cdots \ y_{t|t-1}^{(N)} - \overline{y}_{t|t-1}],$$

$$P_{xy} = \frac{1}{N-1}E_x E_y^T, \quad P_{yy} = \frac{1}{N-1}E_y E_y^T, \quad P_{xx} = \frac{1}{N-1}E_x E_x^T$$

where R is the covariance of ε.

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

$$\begin{aligned} K_t &= P_{xy}(P_{yy} + R)^{-1} \\ x_{t|t}^{(n)} &= x_{t|t-1}^{(n)} + K_t(y_t + \epsilon^{(n)} - y_{t|t-1}^{(n)}) \end{aligned}$$

where $\epsilon^{(n)} \overset{\mathrm{d}}{=} \varepsilon$.

Hedging: The Monte Carlo local optimal hedging weight is given by

$$\theta_t^{\mathbb{Q}} = \frac{P_{xy}}{P_{xx}}$$

and repeat the propagation and updating steps until $t = T$.

5.2 A simulation study on model calibration with additional latent state spot price process

In this section, a model calibration technique using EnKF is presented to test whether there is some effectiveness on parameter estimation with additional latent state, which is known as spot price dynamic process.

The main purpose of using simulation is that with simulation one get a lot of observations very quickly, whereas observing the market only

corresponds to one path of the time series. The drawback is that we have to make the assumption that the stock market follows a certain model.

Data have been simulated from the Bates models, cf. subsection 2. The initial parameters were chosen according to Lindström et al.(2008) to mimic S&P 500 index as closely as possible, see table 4. The asset trajectories are simulated with discrete time steps. The pseudo algorithm of simulation for different models can be found in Section 2.

S_0	V_0	κ	ξ	σ_v	ρ	λ	μ_J	δ_J
1000	0.0576	2.03	0.04	0.38	-0.7	0.59	-0.05	0.07

Table 4: Parameters for the Bates models used in the simulation study.

The total length of the data set is 200 trading days. 100 European call options over a range of 20 strikes with moneyness between 0.65 and 1.35 and 5 maturities was generated each trading day. Our options data has bid-ask spread of approximately \$1, so we have added noise to the option prices and spot prices 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. $\mathbf{R} = \text{cov}(\epsilon) = 1/16$.

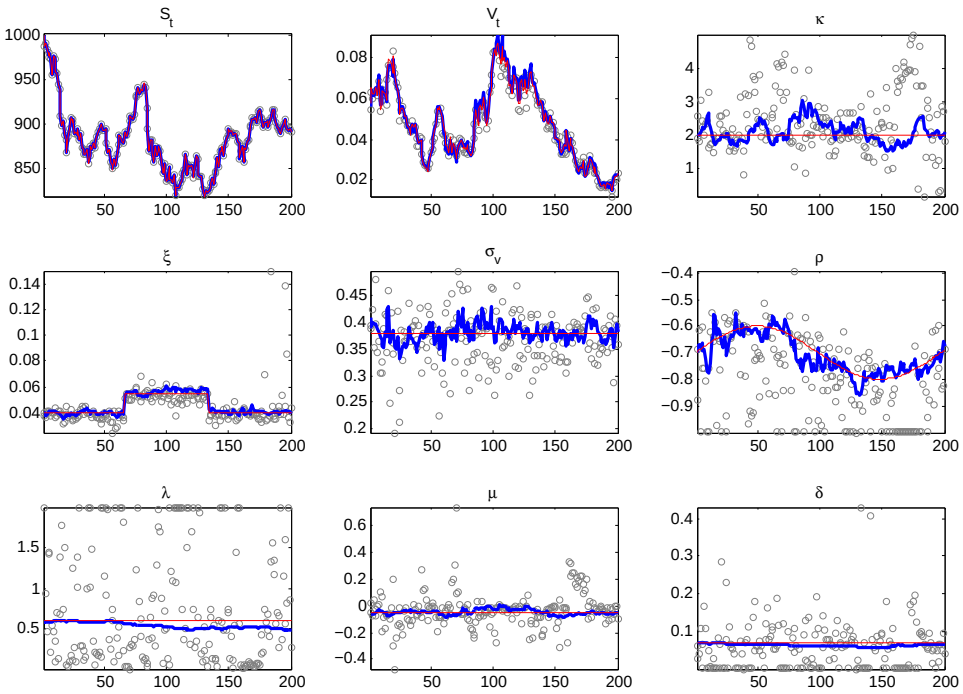

Figure 3: Bates on simulated data. True parameters (thin lines), WLS estimates (circles) and EnKF estimates (thick lines). ξ and ρ are changing over time.

The Bates model has seven parameters and two latent states which are grouped together in the augmented latent state vector, $x = (S_t, V_t, \kappa, \xi, \sigma, \rho, \lambda, \mu, \delta)$. The EnKF and WLS estimated parameters are shown in Figure 3 in addition to the true parameter values. It can be seen that the EnKF estimates are closer to the true parameter values than the WLS estimates. Obviously the filter methods outperform the WLS in terms of parameter estimation.

5.3 An empirical study on simultaneous calibration and hedging

The Bates model and NIGCIR model have been estimated on same market data from options on the S&P 500 index as in last chapter. In order to get the hedging weights for all the options at the same time, the method is applied on the whole data set. The parameter estimates and the calibrated latent states of the Bates model with different ensemble size, are presented in Figure 4, whereas the estimations of the NIGCIR model are shown in Figure 5.

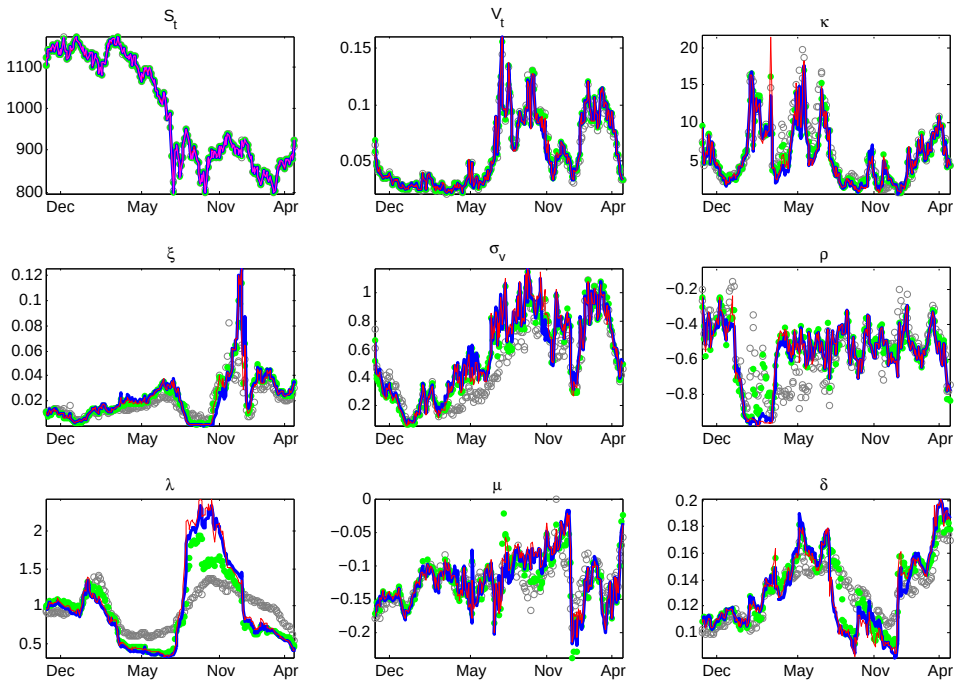

Figure 4: Bates on market data. 50 ensembles (circles), 100 ensembles (dot), 400 ensembles (thick line), and 1000 ensembles (thin line).

Substantially, both figures indicate that this filter technique tracks the spot price process very well. It can be seen from Figure 4 and Figure 5 that parameter estimations show the similar patterns compare to Figure 1 and Figure 2, respectively. It can also be concluded from the figures that the estimations are going to be stable as increasing the number of ensembles for both models. The estimations from EnKF with 400 and 1000 ensembles display similar trajectories for all parameters and latent processes exclusive the affect of randomness for the Bates model. The same situation can be seen from Figure 5 for the NIGCIR model for the volatility parameters. There is a difference for the jump parameters as the spot process plunged. Thus, the NIGCIR model is not the proper model to estimate the spot process for this data set. By using the daily calibration method, different ensemble size may give different but almost equally well-fitted parameters for different model respectively(Table 5).

For the jump parameters, the trajectories change with different number of ensembles. The behavior of jump parameters is way hard to explain.

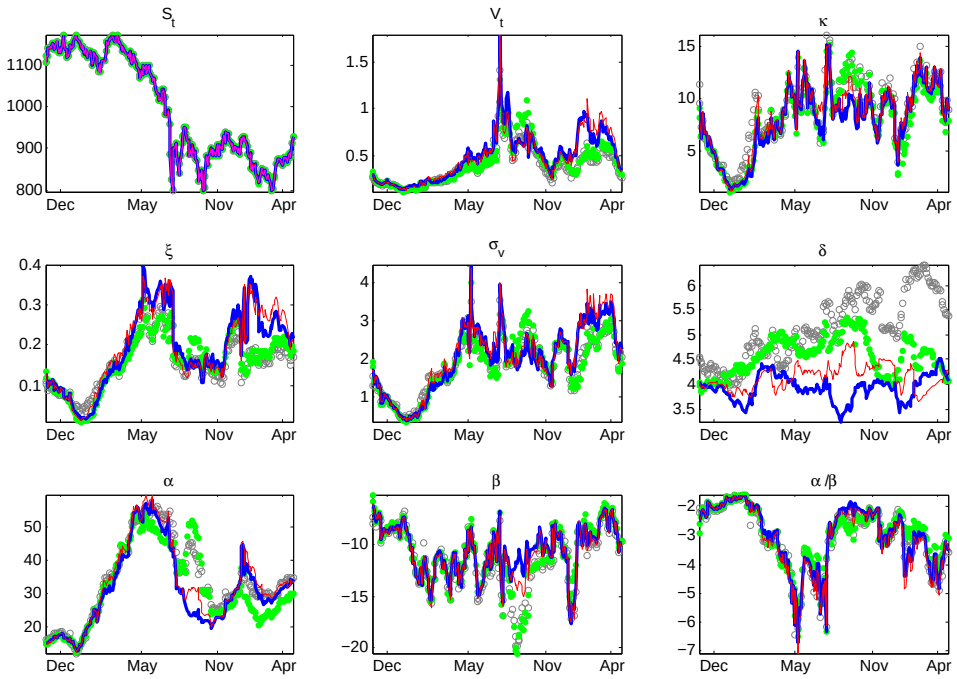

Figure 5: NIGCIR on market data. 50 ensembles (circles), 100 ensembles (dot), 400 ensembles (thick line), and 1000 ensembles (thin line).

The proportional option prices inside spread(IS) and RMSE of the predict prices of global fit are presented in Table 5. We see that adding the underlying asset dynamic to the model causes a minor decline in the level of the proportion of options price inside the bid-ask spread(IS) compared to the measurements in Table 2. The difference of values of IS is caused by the randomness, thus the change of level of IS does not appear to be obvious as the number of ensembles increased.

For the Monte Carlo quadratic hedging strategy, we implement two strategies. Strategy 1 is about to hedge the contract with its underlying asset and risk free asset whereas Strategy 2 is to hedge the contract with its underlying asset and other contracts, which have been explained in Section 2.3. The hedging results are shown in Table 6 for both models with different ensemble size.

The statistics of the hedging error in the above table represent that Strategy 2 performs much better than Strategy 1. Since the value of hedging error from Strategy 2 is almost or even smaller than half of the value of error from Strategy 1. The hedging results show that the value of hedging error decreases with the increasing of the ensemble size. Compare with the result of Strategy 1, it can be obviously seen that Bates model performs better than the NIGCIR model. Since the value of hedging error of Bates model are all less than that of NIGCIR model

with different ensemble size. This is evidence that Bates model is a better model to estimate the spot process for this data set.

It is worth noting that the hedging results we get is for free, i.e. with no additional consumptions except model calibration, since the expensive computation of option values has been already done in the prediction step during model calibration. And it is necessary to mention that errors from both these two Monte Carlo hedging strategy are less than that from the Delta hedge.

	Bates		NIG-CIR	
	IS (%)	RMSE	IS (%)	RMSE
50 ensembles	77.97	0.3205	78.54	0.3010
100 ensembles	80.71	0.2694	76.81	0.3196
400 ensembles	80.03	0.2752	76.80	0.3188
1000 ensembles	79.57	0.2887	76.85	0.3185

Table 5: Summary statistics for the calibration results of the Bates model and the NIGCIR model using EnKF with different ensemble size. The reported numbers are the global fit proportional option prices inside the bid-ask spread(IS%) and error measurements for both predicted options and spot price of S&P 500 data set.

	Bates		NIGCIR	
	Strategy 1	Strategy 2	Strategy 1	Strategy 2
50 ensembles	\$5.94(4.53)	\$3.41(3.50)	\$6.32(4.74)	\$3.14(2.91)
100 ensembles	\$5.85(4.46)	\$2.75(3.22)	\$6.29(4.63)	\$2.79(3.29)
400 ensembles	\$5.63(4.21)	\$2.40(2.65)	\$6.02(4.61)	\$2.37(2.52)
1000 ensembles	**\$5.55(4.10)**	**\$2.12(2.30)**	**\$5.96(4.60)**	**\$2.29(2.31)**

Table 6: The hedging results for the Bates model and the NIGCIR model of two different hedging strategies. Strategy 1 indicate the strategy described in section 2.3.1, named Monte Carlo quadratic hedging with underlying asset and risk-free asset while Strategy 2 is known as Monte Carlo quadratic hedging with additional instrument(section 2.3.2). The reported numbers are the values of average absolute hedging error over the whole period and the standard deviation(in braces).

6 Conclusion

This thesis evaluated the Ensemble Kalman filter technique on model calibration with simulated and market option data by applying two commonly used models with stochastic volatilities and with/without jumps in both underlying asset and volatility dynamics: the Bates model and the NIG-CIR model. The Fourier transform based technique was applied to obtain closed form expressions for the prices of European options.

An empirical study of model calibration was performed on S&P 500 option data. The daily model parameters were calibrated by using the Ensemble Kalman Filter. The estimations have been compared to the standard Weighted Least Square estimates as well as to the Unscented Kalman filter estimates. The overall impression of model calibration is that filter methods provide more stable and more robust estimation than WLS. The UKF and EnKF present the similar results whereas the WLS estimates are not stable, especially for the NIG-CIR model. We find WLS is prone to over-fitting the data and discover the variation of ensemble size does not influence the calibration results. To further study the relationship between ensemble size and calibration accuracy, two hedging strategy known as the Delta hedge and the Monte Carlo Quadratic hedging, were used to test the ability of forecast of models with parameters which are estimated with different number of ensembles. The statistics of values of hedging error also illustrate the affects on predictability of models and that the efficiency of model calibration doesn't appear with variation of ensembles size. However, the level of hedging error goes down with the increasing number of simulations in the Monte Carlo hedging. The number of simulations has the same meaning as ensemble size in the method of simultaneous calibration and hedging. It is not our popurse to outcrop the best ensemble size but enlighten an tendency insight.

Afterwards, the research focused on analyzing the efficiency of simultaneous calibration and hedging. A simulation study first applied on model calibration with additional latent state, named spot price process dynamics. This is the foundation of implementation of the method of simultaneous calibration and hedging. The results show that EnKF works successfully on this. Then we make use of the method of simultaneous calibration and hedging on S&P 500 option data. After using two hedging strategies with four groups of parameters, estimated with different ensemble size, the data showed the two Monte Carlo hedging strategies perform better than the Delta hedge.

Since the Black-Scholes model, the Merton model and the Heston model are sub-models of the Bates model, it is easy to apply the methods on these models.

It should be emphasized that these conclusions have been reached only on the basis of calibrating and hedging S&P 500 index options, especially this data set. However, we believe that a similar behavior can be expected for other datasets.

References

B.D.O. Anderson and J.B. Moore. *Optimal filtering*, volume 11. Prentice-hall englewood cliffs, NJ, 1979.

D.S. Bates. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *Review of financial studies*, 9(1): 69–107, 1996.

F. Black and M. Scholes. The pricing of options and corporate liabilities. *The journal of political economy*, pages 637–654, 1973.

G. Burgers, P. Jan van Leeuwen, and G. Evensen. Analysis scheme in the ensemble kalman filter. *Monthly weather review*, 126(6):1719–1724, 1998.

P. Carr and D. Madan. Option valuation using the fast fourier transform. *Journal of computational finance*, 2(4):61–73, 1999.

P. Carr and L. Wun. The finite moment log stable process and option pricing. *The journal of finance*, 58(2), 2003.

R. Cont and P. Tankov. *Financial modellng with jump processes*. Chapman and Hall, London, 2004.

R. Daley. *Atmospheric data analysis*, volume 2. Cambridge university press, 1993.

E. Derman and I. Kani. Riding on a smile. *Risk*, 7(2):32–39, 1994.

L. Devroye and L. Devroye. *Non-uniform random variate generation*, volume 4. Springer-verlag new york, 1986.

B. Dupire. Pricing with a smile. *Risk*, 7(1):18–20, 1994.

G Evensen. Sequential data assimilation with nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Journal of geophysical research-oceans*, 99(C5):10143–10162, 1994.

G. Evensen. Advanced data assimilation for strongly nonlinear dynamics. *Monthly weather review*, 125(6):1342–1354, 1997.

G. Evensen. Sequential data assimilation for nonlinear dynamics: the ensemble kalman filter. *Ocean forecasting: conceptual basis and applications*, 2002.

H. Foellmer and D Sondermann. Hedging of non-redundant contingent claims. *Contributions to mathematical economics*, pages 205–223, 1986.

H. Föllmer and M. Schweizer. Hedging by sequential regression: An introduction to the mathematics of option trading. *Astin bulletin*, 18 (2):147–160, 1989.

S. Gillijns, O.B. Mendoza, J. Chandrasekar, BLR De Moor, DS Bernstein, and A. Ridley. What is the ensemble kalman filter and how well does it work? In *American control conference, 2006*, pages 6–pp. IEEE, 2006.

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

A. Hommels, F. Molenkamp, AW Heemink, and B. Nguyen. Inverse analysis of an embankment on soft clay using the ensemble kalman filter. In *Proc. of the 10th Int. Conf. on Civil, Structural and Env. Eng. Computing*, volume 252. Civil-Comp Press, Stirling, United Kingdom, paper, 2005.

J. Hull. *Options, futures and other derivatives*. Pearson prentice hall, 2009.

S.J. Julier and J.K. Uhlmann. A new extension of the kalman filter to nonlinear systems. In *Int. Symp. Aerospace/Defense Sensing, Simul. and Controls*, volume 3, page 26. Spie Bellingham, WA, 1997.

R.E. Kalman et al. A new approach to linear filtering and prediction problems. *Journal of basic engineering*, 82(1):35–45, 1960.

E. Kalnay. *Atmospheric modeling, data assimilation and predictability*. Cambridge university press, 2002.

E. Lindström, J. Ströjby, M. Brodén, M. Wiktorsson, and J. Holst. Sequential calibration of options. *Computational Statistics & Data Analysis*, 52(6):2877–2891, 2008.

R.C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of financial economics*, 3(1):125–144, 1976.

P. Samuelson. Rationale theory of warrant pricing. *Industrial management review*, 6, 1965.

W. Schoutens. *Lévy processes in Finance*. Wiley, 2003.

W. Schoutens, E. Simons, and J. Tistaert. A perfect calibration! now what. *Wilmott magazine*, 2:66–78, 2004.

M. Schweizer. Option hedging for semimartingales. *Stochastic processes and their applications*, 37(2):339–363, 1991.