

MAXIMUM LIKELIHOOD ESTIMATION WITH DYNAMIC MEASUREMENT ERRORS AND APPLICATION TO INTEREST RATE MODELING

Masnita Misiran

Department of Mathematics and Statistics
School of Quantitative Sciences
University Utara Malaysia
Kedah, 06010, MALAYSIA

Abstract: Stochastic volatility (SV) model is widely applied in the extension of the constant volatility in Black-Scholes option pricing. In this paper, we extend the SV model driven by fractional Brownian motion (FBM). A crucial problem in its application is how the unknown parameters in the model are to be estimated. We propose the innovation algorithm, and follow by the maximum likelihood estimation approach, which enables us to derive the estimators of parameters involved in this model. We will also present the simulation outcomes to illustrate the efficiency and reliability of the proposed method.

Key Words: stochastic volatility, fractional Brownian motion, innovation algorithm, maximum likelihood estimation

1. Introduction

Black-Scholes model is a jewel in financial world. Ever since its debut in 1973 by Black, extensive works have been conducted to improve the performance of this model. Interested readers can refer to [Elliott and Kopp (2005)] for more details in this study. Recent works implore the possibility of extending the constant volatility in Black-Scholes market to the stochastic volatility (SV). This idea is

Received: May 15, 2016

Revised: September 7, 2016

Published: November 5, 2016

© 2016 Academic Publications, Ltd.

url: www.acadpubl.eu

more reasonable in real application. It allows us to capture the impact of time-varying volatility in financial data. However, many econometrician struggled to estimate these models in early years. Only in 1990s that the estimation and the simulation works for these models begin to develop. Readers are referred to some works in [Melino and Turnbull (1990); Andersen and Sørensen (1996); Harvey *et. al.* (1994)].

This problem keep intrigue researchers. The second generation of works on SV begin to focus on some more advance extension. Some considered jumps in the dynamics of asset prices [Creeel & Kristensen (2015); Cheang, Chiarella & Ziogas (2013); Anderson (2007); Eraker *et. al.* (2003)]. Others considered the long memory property in SV model [Shirota, Hizu & Omori (2014); Bos, Koopman & Ooms (2014); Breidt *et. al.*(1998); Comte and Renault (1998); Andersen and Bollerslev (1997)]. Recent works on this problem can be found in [Casas and Gao (2008); Hu (2004)]. Most of these works assumed the model to be generated by the finite parameter model. They used spectral density to extract the memory parameter in its autocovariances to define the long memory in SV model. Some exceptions are works by Høg and Frederiksen (2006) and Hult (2003). They consider the model in continuous time. Fractional Brownian motion (FBM) is considered to define the long memory process. Høg used Kalman filter approach to estimate the parameters, but state the noise as a stationary ARFIMA $(0, d, 0)$ process. Hult (2003) however, used the spectral density in his estimation process.

The study of FBM as the governing noises in some financial model begin in 1968. This is due to works by Mandelbrot and his colleagues [Mandelbrot & van Ness (1968); Mandelbrot & Wallis (1969); Mandelbrot (1972)]. Although some has addressed FBM in finance in sceptical view [Rogers (1997)], this issue never fail to capture researchers' interest. Work by Hu and Øksendal (2003) has motivated renewed interest and faith to FBM in financial modeling. Readers are referred to works by [Mandelbrot (2004); Taleb (2007)] for a critical view on the randomness by standard Brownian motion (BM). It is reasonable to consider FBM in the SV model. A crucial problem with the real applications of this model, is how to get the unknown values of the parameters in SV models, in particular, the a, σ and H . They represent the drift, diffusion coefficient and Hurst parameter, respectively. Once these parameters are estimated, the valuing process in option can be done in the SV setting.

Most of earlier works consider Ornstein-Uhlenbeck (OU) process to model this SV model. It is natural to extend this standard OU process in the nature of fractality, by substituting FBM to the standard BM. This is called fractional Ornstein-Uhlenbeck (FOU) process. Some works on FOU have appeared in

literature in recent years [Cheridito *et. al.* (2003); Kleptsyna and Le Breton (2002); Høg and Frederiksen (2006)], but they only developed the theoretical investigation on this matter. Some works on FOU that begin to provide estimations on the proposed methods include Brouste & Iacus (2013; Tanaka (2013) and Salomon & Fort (2013), to mention just a few. We noted that Hult (2003) has estimated the parameter involved in FOU, based on discrete observation by using the spectral representation. We would also like to mention work by Hu (2004) who explore this process in financial environment. He considered this process in option pricing problem. However, this exploratory study too, concerned on the theoretical investigation. In this paper, our aim is to use the innovation algorithm, couple with the maximum likelihood (ML) estimation procedure, to give efficient estimate to the parameters involved in this process. We will also present the numerical investigation to better suit the real-life problem.

In what follows, we summarize some important results reported in Ahmed & Charalambous (2002). Let (Ω, \mathcal{F}, P) be a probability space and $H \in (0, 1)$, where H is the index of the self-similarity known as *Hurst parameter*. A random process $\{B_H, t \geq 0\}$, defined on the probability space $(\Omega, \mathcal{F}_e, P)$ is said to be a FBM if the following conditions are satisfied.

- (i) $P\{B_H(0) = 0\} = 1$;
- (ii) for each $t \in \mathbb{R}_+ \equiv [0, \infty)$, $B_H(t)$ is an \mathcal{F}_e -measurable random variable having Gaussian distribution with $E\{B_H(t)\} = 0$;
- (iii) for $t, s \in \mathbb{R}_+$, $E\{B_H(t)B_H(s)\} = \frac{1}{2}\{t^{2H} + s^{2H} - |t - s|^{2H}\}$;
- (iv) for $H > \frac{1}{2}$, the sample paths of B_H are continuous with probability one but nowhere differentiable; and
- (v) B_H is self similar in the sense that, for any $\alpha > 0$, the probability laws of $\{B_H(\alpha t)\}$ and $\{\alpha^H B_H(t)\}$ coincide.

2. Model Simplification

The fractional Ornstein-Uhlenbeck process can be written in the form of

$$\begin{aligned} dx(t) &= -a_1 x(t)dt + \sigma_1 dB_{H_1}(t) \\ x(0) &= x_0, \end{aligned} \tag{1}$$

where $a_1 > 0$ and σ_1 are constant parameters, normally known as drift and diffusion coefficients, respectively. $(B_H(t), t \geq 0)$ is a fractional Brownian motion on some probability space (Ω, \mathcal{F}, P) with *Hurst parameter*, H . This model is used to replace the constant volatility in risky asset model. Some used this model as a short rate model [De Rossi, 2004].

Consider the measurement dynamics as

$$\begin{aligned} dy(t) &= a_2 x(t) dt + \sigma_2 dB_{H_2}(t) \\ y(0) &= 0, \end{aligned} \quad (2)$$

where $a_2 \in \mathbb{R}$ and $\sigma_2 \in \mathbb{R}$ are drift and diffusion coefficients, respectively. In order to simplify this problem, let the systems be considered in discrete processes. The system now becomes:

$$x(t + \Delta t) - x(t) = -a_1 x(t) \Delta t + \sigma_1 (B_{H_1}(t + \Delta t) - B_{H_1}(t)) \quad (3)$$

and the measurement dynamics is

$$y(t + \Delta t) - y(t) = a_2 x(t) \Delta t + \sigma_2 (B_{H_2}(t + \Delta t) - B_{H_2}(t)) \quad (4)$$

By considering $t = k\Delta t$,

$$x((k+1)\Delta t) - x(k\Delta t) = -a_1 x(k\Delta t) \Delta t + \sigma_1 (B_{H_1}((k+1)\Delta t) - B_{H_1}(k\Delta t)) \quad (5)$$

and

$$y((k+1)\Delta t) - y(k\Delta t) = a_2 x(k\Delta t) \Delta t + \sigma_2 (B_{H_2}((k+1)\Delta t) - B_{H_2}(k\Delta t)) \quad (6)$$

Denote $\tilde{X}_{k+1} = x((k+1)\Delta t)$, $\tilde{Y}_k = y((k+1)\Delta t) - y(k\Delta t)$, $\xi_{k+1} = B_{H_1}((k+1)\Delta t) - B_{H_1}(k\Delta t)$, and $\eta_k = B_{H_2}((k+1)\Delta t) - B_{H_2}(k\Delta t)$. Our simplified version now becomes:

$$\tilde{X}_{k+1} = (1 - a_1 \Delta t) \tilde{X}_k + \sigma_1 \xi_{k+1} \quad (7)$$

$$\tilde{Y}_k = a_2 \Delta t \tilde{X}_k + \sigma_2 \eta_k, \quad (8)$$

with $\eta_k \sim N(0, (\Delta t)^{2H_2})$ and $\xi_k \sim N(0, (\Delta t)^{2H_1})$. Note that (7) is an autoregressive (AR) process. We follow iteration process and the Cauchy criterion in [Brockwell and Davis, 1987] and further stated the (AR) process as:

$$\tilde{X}_{k+1} = \sum_{j=0}^{\infty} (1 - a_1 \Delta t)^j \sigma_1 \xi_{k+1-j}, \quad (9)$$

with $|1 - a_1 \Delta t| < 1$.

By using the property on the increment of the FBM, $\{B_H(t + \Delta t) - B_H(t)\}$, the covariances of η and ξ can be expressed as:

$$\gamma_\eta(m) = \text{cov}(\eta_k, \eta_{k-m}) = \frac{1}{2} [| (m+1)\Delta t |^{2H_2} + | (m-1)\Delta t |^{2H_2} - 2|m\Delta t|^{2H_2}] \quad (10)$$

and

$$\gamma_\xi(m) = \text{cov}(\xi_k, \xi_{k-m}) = \frac{1}{2} [| (m+1)\Delta t |^{2H_1} + | (m-1)\Delta t |^{2H_1} - 2|m\Delta t|^{2H_1}], \quad (11)$$

respectively. Note that when $H < \frac{1}{2}$, the increments are negatively correlated whereas $H > \frac{1}{2}$ shows the positive correlation. The increment is a stationary process, which is often referred to as fractional Gaussian noise.

Note that from (7) the covariance for \tilde{X} can be obtained as follows:

$$\begin{aligned} \gamma_{\tilde{X}}(m) &= \text{cov}(\tilde{X}_k, \tilde{X}_{k-m}) \\ &= \text{cov}\left(\sum_{j=0}^{\infty} (1 - a_1 \Delta t)^j \sigma_1 \xi_{k-j}, \sum_{i=0}^{\infty} (1 - a_1 \Delta t)^i \sigma_1 \xi_{k-m-i}\right) \\ &= \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} (1 - a_1 \Delta t)^{j+i} \sigma_1^2 \text{cov}(\xi_{k-j}, \xi_{k-m-i}) \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} (1 - a_1 \Delta t)^{i+j} \sigma_1^2 \gamma_\xi(m + i - j) \end{aligned}$$

By choosing sufficiently large L , (12) can now be written as

$$\gamma_{\tilde{X}}(m) \approx \sum_{i=0}^L \sum_{j=0}^L (1 - a_1 \Delta t)^{i+j} \sigma_1^2 \gamma_\xi(m + i - j) \quad (12)$$

with $|a_1 \Delta t| < 1$.

Further, denote $(\tilde{Y}_1, \dots, \tilde{Y}_T) \sim N((\mu, \dots, \mu), \Sigma)$, with $\Sigma = (\gamma_{\tilde{Y}}(i - j))_{\{i, j=1\}}^T$. By using (8), write covariance for \tilde{Y} as:

$$\begin{aligned} \gamma_{\tilde{Y}}(m) &= \text{cov}(\tilde{Y}_k, \tilde{Y}_{k-m}) \\ &= (a_2 \Delta t)^2 \text{cov}(\tilde{X}_k, \tilde{X}_{k-m}) + (\sigma_2)^2 \text{cov}(\eta_k, \eta_{k-m}) \\ &= (a_2 \Delta t)^2 \gamma_{\tilde{X}}(m) + \sigma_2^2 \gamma_\eta(m). \end{aligned}$$

3. Methodology of Estimation

In this section, our aim is to estimate the parameters $a_1, a_2, \sigma_1, \sigma_2, H_1$ and H_2 involved in the model considered. Here, the likelihood function is presented for our problem.

By considering (8), we illustrate the likelihood function for $(\tilde{Y}_1, \dots, \tilde{Y}_T)$ as follows:

$$L(a_1, a_2, \sigma_1, \sigma_2, H_1, H_2) = \frac{1}{(2\pi)^{T/2}(\det(\Sigma_T))^{1/2}} \exp\left\{-\frac{1}{2}(\tilde{Y} - \mu)' \Sigma_T^{-1}(\tilde{Y} - \mu)\right\} \quad (13)$$

It is known that the maximum likelihood imply the efficient estimates of parameters. However, in (13), it is very difficult to find the gradient of the likelihood function with respect to the parameters. This system is too involved. In order to cater this problem, the innovation algorithm is implemented.

Suppose that $\{\tilde{Y}_1, \dots, \tilde{Y}_T\}$ is stationary and

$$\tilde{Y}_{n|n-1} = \phi_{n1}(\tilde{Y}_n - \mu) + \phi_{n2}(\tilde{Y}_{n-1} - \mu) + \dots + \phi_{nn}(\tilde{Y}_1 - \mu) + \mu \quad (14)$$

The last term of (14) is to ensure $E[\tilde{Y}_n - \tilde{Y}_{n|n-1}] = 0$.

Based on Shumway & Stoffer (2006), the coefficients $\{\phi_{n1}, \phi_{n2}, \dots, \phi_{nn}\}$ satisfy

$$E[(\tilde{Y}_{n+1} - \sum_{k=1}^n \phi_{nk} \tilde{Y}_{n+1-k}) \tilde{Y}_{n+1-k}] = 0, \quad k = 1, \dots, n,$$

or

$$\sum_{k=1}^n \phi_{nk} \gamma_{\tilde{Y}}(k-j) = \gamma_{\tilde{Y}}(k), \quad k = 1, \dots, n. \quad (15)$$

(15) can be rewritten in matrix notation as obtained from the Yule-Walker equations,

$$\mathbf{\Gamma}_n \phi_n = \gamma_n, \quad (16)$$

where $\mathbf{\Gamma}_n = \{\gamma_{\tilde{Y}}(k-j)\}_{j,k=1}^n$ is an $n \times n$ covariance matrix, $\phi_n = (\phi_{n1}, \dots, \phi_{nn})^T$ is an $n \times 1$ vector, and

$$\gamma_n = (\gamma_{\tilde{Y}}(1), \dots, \gamma_{\tilde{Y}}(n))^T \quad (17)$$

is an $n \times 1$ vector. Equation (14) and (16) are known as the one-step prediction equation.

Thus,

$$\phi_n = \mathbf{\Gamma}_n^{-1} \gamma_n. \quad (18)$$

Let $\tilde{\gamma}_{n-1} = (\gamma_{\tilde{Y}}(n-1), \dots, \gamma_{\tilde{Y}}(1))^T$. Then,

$$\mathbf{\Gamma}_n = \begin{bmatrix} \mathbf{\Gamma}_{n-1} & \tilde{\gamma}_{n-1} \\ \tilde{\gamma}_{n-1}^T & \gamma_{\tilde{Y}}(0) \end{bmatrix}.$$

Thus

$$\mathbf{\Gamma}_n^{-1} = \begin{bmatrix} I & -\mathbf{\Gamma}_{n-1}^{-1}\tilde{\gamma}_{n-1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{\Gamma}_{n-1}^{-1} & 0 \\ 0 & (\gamma_{\tilde{Y}}(0) - \tilde{\gamma}_{n-1}^T \mathbf{\Gamma}_{n-1}^{-1} \tilde{\gamma}_{n-1})^{-1} \end{bmatrix} \times \begin{bmatrix} I & 0 \\ -\tilde{\gamma}_{n-1}^T \mathbf{\Gamma}_{n-1}^{-1} & 1 \end{bmatrix}. \quad (19)$$

According to (20), we can compute $\mathbf{\Gamma}_n^{-1}$ recursively. Let

$$\begin{aligned} \varepsilon_n &= \tilde{Y}_n - \tilde{Y}_{n|n-1} \\ &= \tilde{Y}_n - \mu - \sum_{k=1}^n \phi_{nk}(\tilde{Y}_k - \mu). \end{aligned}$$

where ε_n is a sequence of random variables with mean zero and standard deviation v_n such that $\{v_n^{-1}\varepsilon_n\}$ is i.i.d, ϕ_{nk} and v_n are respectively the autoregressive and standard deviation parameters.

Thus,

$$\begin{aligned} &\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_T \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ -\phi_{11} & 1 & 0 & \dots & 0 \\ \vdots & \dots & \dots & \dots & \dots \\ -\phi_{(T-1)1} & -\phi_{(T-1)2} & \dots & -\phi_{(T-1)(T-1)} & 1 \end{pmatrix} \begin{pmatrix} \tilde{Y}_1 - \mu \\ \tilde{Y}_2 - \mu \\ \vdots \\ \tilde{Y}_T - \mu \end{pmatrix}. \quad (20) \end{aligned}$$

Let $\mathbf{Y} = [\tilde{Y}_1 - \mu, \dots, \tilde{Y}_T - \mu]^T$, and $\boldsymbol{\varepsilon} = [\varepsilon_1, \dots, \varepsilon_T]$, where $\varepsilon_T \sim N(0, v_T^2)$, with v_T^2 is the mean square prediction error,

$$v_T^2 = \gamma(0) - \boldsymbol{\gamma}_T^T \mathbf{\Gamma}_T^{-1} \boldsymbol{\gamma}_T. \quad (21)$$

(20) can be rewritten as

$$\boldsymbol{\varepsilon} = \mathbf{A}\mathbf{Y}.$$

Thus,

$$\mathbf{Y} = \mathbf{A}^{-1}\boldsymbol{\varepsilon}.$$

Hence, the autocovariance function,

$$\begin{aligned}\boldsymbol{\Sigma}_T &= \text{cov}(\mathbf{Y}, \mathbf{Y}) = E(\mathbf{Y}\mathbf{Y}^T) = \mathbf{A}^{-1}E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T](\mathbf{A}^{-1})^T \\ &= \mathbf{A}^{-1} \begin{bmatrix} E(\varepsilon_1)^2 & 0 & \dots & 0 \\ 0 & E(\varepsilon_2)^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & E(\varepsilon_T)^2 \end{bmatrix} (\mathbf{A}^{-1})^T.\end{aligned}$$

From (22), $\boldsymbol{\Sigma}_T^{-1}$ can be written as

$$\boldsymbol{\Sigma}_T^{-1} = \mathbf{A}^T \begin{bmatrix} \frac{1}{v_1^2} & 0 & \dots & 0 \\ 0 & \frac{1}{v_2^2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \frac{1}{v_T^2} \end{bmatrix} \mathbf{A}. \quad (22)$$

and

$$\det(\boldsymbol{\Sigma}_T) = \prod_{i=1}^T E(\varepsilon_i)^2 = \prod_{i=1}^T v_i^2. \quad (23)$$

From (22) and (23), the likelihood function (13) can now be calculated.

Now, the likelihood function is transformed into the following optimization problem.

Problem P. Maximizes the cost function

$$L(\theta) \quad (24)$$

subject to

$$\begin{aligned}E(\tilde{Y} - \mu)^2 &\geq 0 \\ v^2 &\geq 0.\end{aligned} \quad (25)$$

where $\theta = (a_1, a_2, \sigma_1, \sigma_2, H_1, H_2)$.

This optimization problem is very difficult to solve. The constraints are too involved with covariance functions. To simplify this problem, we use the constraint transcription method reported in Jennings and Teo 1990.

Maximizes the cost function:

$$L(\theta) \quad (26)$$

subject to

$$g_i(\theta) \leq 0, \quad i = 1, 2, \quad (27)$$

where g_i are the constraints in the original problem. Let this problem be referred to as Problem P. For each $i = 1, 2$, we approximate g_i with $G_{i,\varepsilon}(\theta)$, where

$$G_{i,\varepsilon}(\theta) = \begin{cases} g_i, & g_i > \varepsilon \\ \frac{(g_i + \varepsilon)^2}{4\varepsilon}, & -\varepsilon < g_i < \varepsilon \\ 0, & g_i < -\varepsilon, \end{cases} \quad (28)$$

where ε some small number. We now append the approximate functions $G_{i,\varepsilon}$ into the cost function $L(\theta)$ to an appended cost function given below.

Problem $P_{\varepsilon,\gamma}$

$$\hat{L}(\theta) = -L(\theta) - \gamma \sum_{j=1}^m G_{j,\varepsilon}(\theta), \quad (29)$$

where $\gamma > 0$ is a penalty parameter. This is an unconstraint optimization problem, which is referred to as Problem $P_{\varepsilon,\gamma}$. It is known (see Jenning and Teo 1990) that for any given $\varepsilon > 0$, there exists a $\gamma(\varepsilon)$ such that, for $\gamma > \gamma(\varepsilon)$, the solution of Problem $P_{\varepsilon,\gamma}$ will satisfy the constraints of Problem P. Let $\hat{\gamma}(\varepsilon)$ be such a γ for each $\varepsilon > 0$. Furthermore, the solution of Problem $P_{\varepsilon,\hat{\gamma}(\varepsilon)}$ converges to the solution of Problem P.

We propose an algorithm to solve Problem $P_{\varepsilon,\gamma}$.

3.1. Algorithm

Here, we present an algorithm to compute the likelihood function using the innovation algorithm.

Step 0. Set the initial value of parameters, $\{a_1, a_2, \sigma_1, \sigma_2, H_1, H_2\}$. Initialize $\gamma_\xi(0) = 0$, $\gamma_\eta = 0$, $\gamma_{\tilde{X}}(0) = 0$ and $\gamma_{\tilde{Y}} = 0$ and set L to be a large number.

Step 1. By updating the functions, calculate $\gamma_\eta(m)$, $\gamma_\xi(m)$, $\gamma_{\tilde{X}}(m)$ and $\gamma_{(\tilde{Y}-\mu)}(m)$, where m from 1 to L .

Step 2. Set k from 2 to T . Compute Γ_k^{-1} in (19).

Step 3. Compute ϕ_k in (18) and v_k in (21).

Step 4. Calculate Σ_T^{-1} and $\det \Sigma_T$ in (22) and (23), respectively.

Step 5. Calculate the updated likelihood function, $L_T(a_1^T, a_2^T, \sigma_1^T, \sigma_2^T, H_1^T, H_2^T)$. This provides the efficient parameters, $\{a_1^T, a_2^T, \sigma_1^T, \sigma_2^T, H_1^T, H_2^T\}$ involved in this system.

4. Simulation Study

In order to examine the performance of the proposed estimators, we have carried out some simulation experiments. First, we generate the data from model (8). We take the parameters $a_1 = 2, a_2 = 2.5, \sigma_1 = 1, \sigma_2 = 0.5, H_1 = 0.7$ and $H_2 = 0.6$, while $\Delta t = \frac{1}{5}$. We simulate the time series from this discrete time model and apply our methodology to estimate the parameters $\vartheta = (a_1, a_2, \sigma_1, \sigma_2, H_1, H_2)'$ using the simulated data set. The simulated annealing method is used in order to find the optimal parameters, simultaneously. The simulation is repeated one hundred times. The simulated outcomes of the average value of estimates based on 100 replications, with bias and variances, are reported in Table 1. The 5 cases of sample sizes $n = 100, 200, 300, 400, 500$ are considered in the table.

Table 1: Average value of estimates based on 100 replications, with bias in () and variance in []

n	100	200	300	400	500
a_1	2.4156 (0.4156) [0.5958]	2.4233 (0.4233) [0.4429]	2.2133 (0.2133) [0.2144]	2.2085 (0.2085) [0.2118]	2.0884 (0.0884) [0.0736]
a_2	2.4829 (-0.0171) [0.1137]	2.5533 (0.0533) [0.0695]	2.5046 (0.0046) [0.0631]	2.4917 (-0.0083) [0.0520]	2.4814 (-0.0186) [0.0527]
σ_1	0.9898 (-0.0102) [0.0951]	1.0329 (0.0329) [0.0313]	0.9879 (-0.0122) [0.0069]	0.9981 (-0.0019) [0.0164]	0.9707 (-0.0293) [0.0059]
σ_2	0.5556 (0.0556) [0.0118]	0.5387 (0.0387) [0.0087]	0.5419 (0.0419) [0.0047]	0.5445 (0.0445) [0.0077]	0.5455 (0.0455) [0.0054]
H_1	0.7037 (0.0037) [0.0012]	0.6998 (-0.00025) [0.000345]	0.7046 (0.0046) [0.00046]	0.7026 (0.0026) [0.00065]	0.7062 (0.0062) [0.00036]
H_2	0.6084 (0.0084) [0.0016]	0.6067 (0.0067) [0.000625]	0.6048 (0.0048) [0.000416]	0.6042 (0.0042) [0.00058]	0.6024 (0.0024) [0.00031]

From the results obtained in Table 1, it shows that our methodology is efficient. Most of the biases and variance obtained are within an acceptable tolerance. All of our estimates for σ_1, σ_2, H_1 and H_2 are obviously quite stable and less biased. The performance on the estimation of a_1 and a_2 are fairly satisfactory. This simulation outcomes indicate that our methodology is promising in obtaining statistically efficient estimators for FOU.

5. Empirical Results

5.1. Data

We used a data set from federal reserve interest rate available online at <http://www.federalreserve.gov>. The interest rate of business day from 2 January 2009 to 31 December 2009 is examined, with 252 observations. A summary of the time series can be found in Table 2, where the mean of this series is 0.1597 and the variance is 0.001502.

Table 2: Summary of the interest rate

Min.	1st Qu.	Median	Mean	Var	3rd Qu.	Max.
0.05	0.13	0.16	0.1597	0.0015	0.18	0.25

5.2. Estimation

We present in this subsection the results of our study of modeling the data by FOU. Let us describe how the estimation procedure is conducted. First, we use many different cases of the initial values in the model estimation procedure. Each of these cases is computed by using the simulated annealing method. This method produces optimal parameters for each case of initial values considered. Then, we examine the likelihood values for these cases. The likelihood that is the maximum is chosen as the suggested estimates. In Table 3, we show some results from these cases, where the initial values are

- i : $a_1 = 2$, $a_2 = 2.5$, $\sigma_1 = 1$, $\sigma_2 = 0.5$, $H_1 = 0.7$ and $H_2 = 0.6$;
- ii : $a_1 = 3$, $a_2 = 5$, $\sigma_1 = 0.5$, $\sigma_2 = 0.7$, $H_1 = 0.8$ and $H_2 = 0.6$;
- iii : $a_1 = 0.1597$, $a_2 = 0.1597$, $\sigma_1 = 0.0387$, $\sigma_2 = 0.0387$, $H_1 = 0.8$ and $H_2 = 0.7$;
- iv : $a_1 = 1$, $a_2 = 2$, $\sigma_1 = 1$, $\sigma_2 = 0.7$, $H_1 = 0.7$ and $H_2 = 0.7$;
- v : $a_1 = 3$, $a_2 = 1$, $\sigma_1 = 0.6$, $\sigma_2 = 0.8$, $H_1 = 0.8$ and $H_2 = 0.8$;
- vi : $a_1 = 3$, $a_2 = 5$, $\sigma_1 = 0.5$, $\sigma_2 = 0.7$, $H_1 = 0.5$ and $H_2 = 0.5$;

- vii : $a_1 = 3$, $a_2 = 5$, $\sigma_1 = 0.5$, $\sigma_2 = 0.7$, $H_1 = 0.51$ and $H_2 = 0.51$;
- viii : $a_1 = 1$, $a_2 = 2$, $\sigma_1 = 1$, $\sigma_2 = 0.7$, $H_1 = 0.5$ and $H_2 = 0.5$;
- ix : $a_1 = 5$, $a_2 = 2$, $\sigma_1 = 1$, $\sigma_2 = 0.5$, $H_1 = 0.7$ and $H_2 = 0.7$;
- x : $a_1 = 5$, $a_2 = 3$, $\sigma_1 = 1$, $\sigma_2 = 0.7$, $H_1 = 0.6$ and $H_2 = 0.8$.

Here, we take $\Delta t = \frac{1}{5}$ which is similar to that of previous section. We can see from Table 3 that the suggested estimates are $a_1 = 2.277$, $a_2 = 4.237$, $\sigma_1 = 0.509$, $\sigma_2 = 0.713$, $H_1 = 0.809$ and $H_2 = 0.982$. Obviously the long-memory property is rather strong in interest rate process.

Table 3: Likelihood value of different initials

Initial	\hat{a}_1	\hat{a}_2	$\hat{\sigma}_1$	$\hat{\sigma}_2$	\hat{H}_1	\hat{H}_2	likelihood value
i	4.529	0.557	0.559	0.917	0.699	0.94	-20629.275
ii	1.563	16.11	0.157	0.636	0.773	0.89	27929.3925
iii	0.227	10.167	0.063	4.33	0.865	0.997	-41402.1124
iv	5.999	0.889	0.966	1.073	0.878	0.913	3340.3349
v	0.813	7.112	0.254	3.818	0.711	0.99	-63041.7107
vi	1.601	15.989	0.096	0.585	0.721	0.868	-94979.9325
vii	1.833	15.851	0.081	1.161	0.521	0.9	90338.7164
viii	2.786	1.088	2.681	0.042	0.921	0.593	-1189.4721
ix	2.277	4.237	0.509	0.713	0.809	0.982	1900437.9156
x	6.414	1.947	0.545	0.585	0.617	0.846	-45609.6661

6. Conclusion

In this paper, we proposed a novel method for estimating the unknown parameters in fractional Ornstein-Uhlenbeck (FOU) model. The likelihood function for FOU is difficult to solve analytically. Its covariance calculation is very expensive. Due to this, we proposed the innovation algorithm approach to simplify this problems. The likelihood function now transformed to an optimization problem with some constraints. The constraints transcription method is applied to append the approximate constraints to the cost function. Then, we

solved a standard unconstrained optimization problem, by using simulated annealing method. We carried out some simulation study so as to illustrate the efficiency of our method. We also carried out empirical study to the interest rate data.

Acknowledgments

This work was supported by fundamental research grant scheme (SO Code: 11992) from Malaysian Ministry of Higher Education.

The author would like to thank the referee for careful reading of the article.

References

- [1] N.U. Ahmed, C.D. Charalambous, Filtering for linear systems driven by fractional Brownian motion, *Journal Control Optim.*, **41**, No. 1 (2002), 313-330.
- [2] T.G. Andersen, T. Bollerslev, Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns, *Journal of Finance*, **52**, (1997), 975-1005.
- [3] T.G. Andersen, T. Bollerslev, F.X. Diebold, C. Ega, Real-time price discovery in global stock, bond and foreign exchange markets, *Journal of International Economics*, **73** (2007), 251-277.
- [4] T.G. Andersen, B. Sørensen, GMM estimation of a stochastic volatility model: A Monte Carlo study, *Journal of Business and Economic Statistics*, **14** (1996), 328-352.
- [5] F.J. Breidt, N. Crato, P. de Lima, On the detection and estimation of long memory in stochastic volatility, *Journal of Econometrics*, **83** (1998), 325-348.
- [6] A. Brouste, S.M. Iacus, Parameter estimation for the discretely observed fractional Ornstein-Uhlenbeck process and the Yuima R Package, *Computational Statistics*, **28**, No. 4 (2013), 1529-1547.
- [7] C.S. Bos, S.J. Koopman, M. Ooms, Long memory with stochastic variance model: A recursive analysis for us inflation, *Computational Statistics, Data Analysis*, **76** (2014), 144-157.
- [8] P.J. Brockwell, R.A. Davis, *Time Series: Theory and Methods*, Springer-Verlag (1987).
- [9] I. Casas, J. Gao, Econometric estimation in long-range dependent volatility models: Theory and practice, *Journal of Econometrics*, **147** (2008), 72-83.
- [10] G.H. Cheang, C. Chiarella, A. Ziogas, The representation of American option prices under stochastic volatility and jump-diffusion dynamics, *Quantitative Finance*, **13**, No. 2 (2013), 241-253.
- [11] F. Comte, E. Renault, Long memory in continuous-time stochastic volatility models, *Mathematical Finance*, **8** (1998), 291-323.
- [12] M. Creel, D. Kristensen, ABC of SV: Limited information likelihood inference in stochastic volatility jump-diffusion models, *Journal of Empirical Finance*, **31** (2015) 85-108.

- [13] G. De Rossi, Kalman filtering of consistent forward rate curves: A tool to estimate and model dynamically the term structure, *Journal of Empirical Finance*, **11** (2004), 277-308.
- [14] B. Erakar, M. Johannes, N.G. Polson, The impact of jumps in returns and volatility, *Journal of Finance*, **53** (2003), 1269-1300.
- [15] A.C. Harvey, E. Ruiz, N. Shephard, Multivariate stochastic variance models, *Review of Economic Studies*, **61** (1994), 247-264.
- [16] E.P. Hog, P.H. Frederiksen, The fractional Ornstein-Uhlenbeck process: Term structure theory and application, *Scandinavian Working Papers in Business Administration* (2006), No:F-2006-01.
- [17] Y. Hu, Option pricing in a market where the volatility is driven by fractional Brownian motions, *Probability, Finance and Insurance* (2004), 164-173.
- [18] H. Hult, Topics on fractional Brownian motion and regular variation for stochastic processes, *Doctoral Dissertation, Royal Institute of Technology* (2003).
- [19] L.S. Jennings, K.L. Teo, A computational algorithm for functional inequality constrained optimization problems, *Automatica*, **26** (1990), 371-375.
- [20] A. Melino, S.M. Turnbull, Pricing foreign currency options with stochastic volatility, *Journal of Econometrics*, **45** (1990), 239-265.
- [21] L.A. Salomon, J.C. Fort, Estimation of the hurst parameter in some fractional processes, *Estimation of the Hurst Parameter in Some Fractional Processes*, **83**, No. 3 (2013), 542-554.
- [22] S. Shirota, T. Hizu, Y. Omori, Realized stochastic volatility with leverage and long memory, *Computational Statistics, Data Analysis*, **76** (2014), 618-641.
- [23] R.H. Shumway, D.S. Stoffer, *Time Series Analysis and Its Applications*, Springer (2006).
- [24] K. Tanaka, Distributions of the maximum likelihood and minimum contrast estimators associated with the fractional ornstein uhlenbeck process, *Statistical Inference for Stochastic Processes*, **16**, No. 3 (2013), 173-192.