Research Article

Gaussian Estimation of One-Factor Mean Reversion Processes

Freddy H. Marín Sánchez and J. Sebastian Palacio

Basic Science Department, Eafit University, Carrera 49 No. 7 Sur 50, Medellin, Colombia

Correspondence should be addressed to Freddy H. Marín Sánchez; fmarinsa@eafit.edu.co

We propose a new alternative method to estimate the parameters in one-factor mean reversion processes based on the maximum likelihood technique. This approach makes use of Euler-Maruyama scheme to approximate the continuous-time model and build a new process discretized. The closed formulas for the estimators are obtained. Using simulated data series, we compare the results obtained with the results published by other authors.

1. Introduction

Different applications of stochastic differential equations have attracted the attention of many researchers in relation to the theoretical advances needed for modeling and simulation of the dynamic behavior of the main variables affecting the evolution of some phenomena. Applications in fields of economics, finance, physics, insurance, and others have been drivers of these developments. Particularly, through diffusion processes have been carried out modeling interest rates and commodity prices, where they used in their modeling processes mean reversion. Some trends of these studies are based on some a priori knowledge to determine the structure of the model to be used, while other alternatives are part of a general structure, and statistical procedures through the particular structure are determined to represent the phenomenon. Whichever approach is used, in general, in the specification of the model, it is necessary to estimate parameters to achieve a proper fit of the model.

Although some parameter estimation methods have been extensively studied in diffusion processes, this has a serious problem, which is that the specification of the models is done in continuous time but the data that is available for adjusting the parameters are discretely sampled in time. A common solution to this problem is to discretize the continuous-time model based on Euler-Maruyama scheme and perform procedures on the theoretical formulation resulting in discrete time. From this scheme alternatives can be proposed for parameter estimation approach among which stands out the maximum likelihood estimation.

Following the classification proposed by Yu and Phillips [1], different parameter estimation procedures can be classified globally into three categories according to how to proceed from the model formulation: discretization of continuous process and estimation of the parameters with the resulting discretized process, estimation of the transition probability function for the continuous-time model, and independent estimations of the trend and volatility components by kernel methods.

Thus, among the methods that are based on the discretization of the continuous-time model is Nowman [2], who considers a stochastic differential equation with two unknown variables in the drift term and two in the volatility, elasticity allowing the variance. Similarly, Yu and Phillips [1] present an improvement to the method proposed by Nowman [2], which is based on a nonuniform sampling of the observations, and the estimation process makes use of the equivalence of the methods of maximum likelihood and least squares under the conditions laid down in the model. Alternatively, Yu [3] proposes a new scheme that is based on the idea of transforming a stochastic differential equation (EDE) in its discrete-time equivalent, an AR(1) to determine the formulas that approximate the bias error present in the estimation of one of the parameters.

Some of the works that could be classified as methods in which it seeks to find the transition probability function and use Martingale properties to approximate the likelihood function are Durham and Gallant [4], Valdivieso et al. [5], and Elerian et al. [6]. Durham and Gallant [4] propose a scheme based on simulation. Establishing different model modifications, Pedersen [7] improves computational performance, and characteristics of the model are studied using simulated synthetic series. Koulis and Thavaneswaran [8] study combined martingale estimation functions for discretely observed interest rates models. They derive closed-form expressions for the first four conditional moments of the CIR model via Ito's formula and for extended versions of the CIR model using the Milstein's approximation and construct the optimal estimating functions. They also show that combined estimating functions have more information when the conditional mean and variance of the observed process depend on the parameter of interest. Valdivieso et al. [5] proposed a method for estimating the transfer function through the inversion of the characteristic function of the process and the use of Fourier transform. This methodology is developed for some particular cases of D-OU processes (Ornstein-Uhlenbeck processes with broken car distribution D). Elerian et al. [6] used simulation-based techniques for estimating the transition probability depending on the sample to incorporate auxiliary dummy observations. The parameters are obtained by simulation and optimization using numerical algorithms.

Methods based on kernel techniques approximate density functions of the components and spreading tendency of the process. The works of Jiang and Knight [9] and Florens-Zmirou [10] are within this framework.

In this paper, we propose an alternative scheme for parameter estimation in one-factor mean reversion models, using the Euler-Maruyama discretization. Using the properties of discretized error normality, the maximum likelihood technique to obtain closed formulas for the estimators applies. We present some performance measures of the maximum likelihood estimators obtained with the proposed method in simulated processes, and we compare them with the results published by Nowman [2], Yu and Phillips [1], and Durham and Gallant [4].

This paper is organized as follows. In Section 2, a brief description of mean reversion processes of one factor with constant parameters. Some Gaussian estimation techniques are presented in Section 3. Section 4 develops the proposed method, and Section 5 presents some experimental numerical results with different sampling rates.

2. Mean Reversion Processes

Mean reversion processes have been considered as a natural choice to describe the dynamic behavior of interest rates in Vasicek [11], Brennan and Schwartz [12], Cox et al. [13], and Chan et al. [14]; daily prices of some commodities such as oil and natural gas in Lari-Lavassani et al. [15] have a fundamental role in modeling spot prices when they exhibited long-term trends, such as electrical energy (Pilipovic [16]).

Mean reversion processes of one factor with constant parameters can be written as

$$dX_t = \alpha \left(\mu - X_t \right) dt + \sigma X_t^\gamma dB_t; \quad t \in [0, T] \qquad (1)$$

with initial condition $X_0 = x$, where $\alpha > 0$, $\sigma > 0$, $\mu \in \mathbb{R}$, and $\gamma \in [0, 3/2]$ are constants and $\{B_t\}_{t \geq 0}$ is a unidimensional standard Brownian motion defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

The α parameter is called reversion rate, μ is the mean reversion level or trend of long-run equilibrium, σ is the parameter associated with the volatility, and γ determines the sensitivity of the variance to the level of X_t.

The model (1) is known as CKLS proposed by Chan et al. [14] and is a generalization of the models of interest rate in the short term with constant parameters in which $\gamma = 0$, $1/2$, and 1 or $\gamma = 3/2$ with $\alpha = 0$.

In general, mean reversion processes are stationary. For example, note that (1) can be written in integral form as

$$X_t - X_0 = \alpha \int_0^t \left(\mu - X_s \right) ds + \sigma \int_0^t X_s^\gamma dB_s. \qquad (2)$$

Taking the expected value, $E[X_t] - E[X_0] = \alpha \int_0^t (\mu - E[X_s]) ds$ and thus we obtain the differential equation

$$\dot{m}(t) = \alpha \left(\mu - m(t) \right), \quad m(t) = E\left[X_t \right]. \qquad (3)$$

The solution of this equation is $m(t) = \mu + (X_0 - \mu) e^{-\alpha t}$, so $\lim_{t \to \infty} E[X_t] = \lim_{t \to \infty} m(t) = \mu$.

It can be concluded then that $m(t) = \mu$ is the equilibrium level for the linear differential equation. In other words, the expected value of the process $m(t)$ returns in the long term to μ. In a finite-time horizon, μ plays the role of attractor in the sense that when $X_t > \mu$ the trend term $\alpha(\mu - X_t) < 0$, and therefore, X_t decreases and when $X_t < \mu$ a similar argument establishes that X_t grows.

The stationarity condition is the key to perform the Gaussian estimation of parameters α, σ, and μ, assuming that the elasticity ($\gamma = 0, 1/2, 1, 3/2$) is given a priori.

Although the general model (1) derives various interest rate models studied in the literature, our interest is focused primarily on those models that are strictly stationary.

For example, when $\gamma = 0$, it is said that the process has additive noise so that we get the linear stochastic differential equation

$$dX_t = \alpha \left(\mu - X_t \right) dt + \sigma dB_t; \quad X_0 = x, \ t \in [0, T]. \qquad (4)$$

This model is also known as Ornstein-Uhlenbeck mean reversion process type used by Vasicek [11] in modeling the equilibrium price of the coupon. This Gaussian process has been widely used in the valuation of options on bonds, futures, options on futures, and other financial derivative instruments, while in Lari-Lavassani et al. [15] it has been used to study the dynamic behavior of the prices of oil and natural gas.

Although the model of Vasicek [11] is analytically tractable, it has some drawbacks. Since the interest rate in the short term is normally distributed, for each $t \in [0, T]$ exists

Model	α	μ	σ	γ
Vasicek				0
Brennan-Schwartz				1
CIR SR				1/2
CIR VR	0			3/2
CKLS				[0, 3/2]

a probability that X_t is negative, and this is not reasonable from an economic standpoint. Another drawback of this model is that conditional volatility assumes changes in the interest rate that are constants, independent of the level of X_t.

If $\gamma = 1$ it is said that the process has proportional noise and gives the nonhomogeneous linear stochastic differential equation given by

$$dX_t = \alpha \left(\mu - X_t\right) dt + \sigma X_t dB_t; \quad X_0 = x, \, t \in [0, T]. \quad (5)$$

This model is used by Brennan and Schwartz [12] to obtain a numerical model of convertible bond prices. This process is also used by Courtadon [17] in developing a model for option pricing on discount bonds.

When $\gamma = 1/2$, we obtain the nonlinear stochastic differential equation

$$dX_t = \alpha \left(\mu - X_t\right) dt + \sigma X_t^{1/2} dB_t,$$
$$2\alpha\mu \geq \sigma^2, \quad X_0 = x, \quad t \in [0, T]. \quad (6)$$

This process is naturally linked to the χ centered-square distribution and is the (square root) version of the model of Brennan and Schwartz [12], proposed by Cox et al. [18], which produces types of instantaneous short-term interest strictly positive.

This model has also been used extensively in the development of pricing models for derivative instruments sensitive to interest rates. Examples include pricing models for mortgage-backed in Dunn and McConnell [19], models of options for discount bonds in Cox et al. [18], future and option pricing models for futures in Ramaswamy and Sundaresan [20], the swap valuation models in Sundaresan [21], and option pricing models for the convenience yield in Longstaff [22].

Finally, for the special case in which $\gamma = 3/2$ and $\alpha = 0$, equation (1) takes the form

$$dX_t = \sigma X_t^{3/2} dB_t, \quad X_0 = x, \, t \in [0, T]. \quad (7)$$

This model is introduced by Cox et al. [13] in their study of variable-rate (VR) securities. A similar model is also used by Constantinides and Ingersoll [23] for value bonds in the presence of taxes.

Table 1 shows the summary of the models associated with (1) that are of interest in this work (Vasicek, Brennan-Schwartz, CIR-SR, CIR-VR, and CKLS).

3. Gaussian Estimation of Diffusion Models

Many theoretical models in economics and finance are formulated in continuous time as either a diffusion or diffusion systems. For this reason, identification, estimation, and,

therefore, the investigation of the asymptotic properties of samples of the estimators of these processes turn out to be a difficult task. These problems arise mainly due to the lack of availability of continuous sampling observations. In fact, the estimation of diffusion models has been considered in the literature for many years, and in most scientific journals it only refers to the estimate considering continuous sampling observations, although the data that these models described can only be sampled at discrete points in time (See, e.g., Jiang and Knight [9] and references therein.)

There are various techniques and methods for estimation of models of mean reversion of a single factor; with different structural forms in the trend and distribution functions, most simulation techniques-based on statistical and quite rigorous arguments. The work of Jiang and Knight [9] proposes a parametric identification and estimation procedure for diffusion processes based on discrete sampling observations. The nonparametric kernel estimator for diffusion functions developed in this work deals with general diffusion processes and avoids any functional form specification for either the trend function or distribution function. They show that under certain regularity conditions, the nonparametric estimator function is consistent, and timely dissemination also asymptotically follows a mixed normal distribution. In Jiang and Knight [24], there are some numerical experimental results obtained by a Monte Carlo study using the results of Jiang and Knight [9], which include parameter estimation for Ornstein-Uhlenbeck process. In this sense, the lack of analytical solutions has hampered empirical evidence and validation of alternative hypotheses about these continuous-time models.

Although the maximum likelihood method is a widely used parametric method, the analytical expressions for the likelihood functions exist only for processes of mean reversion with additive noise. In fact, in most of the work on estimation of diffusion models, processes become more complex processes of Ornstein-Uhlenbeck type. For example, Giet and Lubrano [25] consider a diffusion process in which the elasticity constant of the nonlinear term of diffusion can be freely chosen, achieving after reducing it a transformation to an Ornstein-Uhlenbeck process. The major drawback is that there is no degree of freedom in the form of the trend function which is completely determined by the shape of the distribution function and the choice of the underlying linear model. Valdivieso et al. [5] proposed a methodology to estimate maximum likelihood parameters of a one-dimensional stationary process of Ornstein-Uhlenbeck type, which is driven by general Levy process and which is constructed through a distribution D self-decomposed. This approach is based on the inversion of the characteristic function and using the discrete fractional fast Fourier transform. The maximization of the likelihood function is performed using numerical methods.

The framework proposed in our work has some features in common with Nowman [2], Yu and Phillips [1], and Elerian et al. [6]. The similarity of our approach lies in the use of Euler numerical scheme, using discretely sampled data and maximum likelihood estimation, but the main difference is that in our proposal does not need any numerical

optimization because it is possible to obtain a closed formula for the estimators.

For example, in Elerian et al. [6] developed a technique based on Bayesian estimation of nonlinear EDEs approximating the Markov transition densities of broadcasts, including Ornstein-Uhlenbeck processes and elasticity models for interest rates, when observations are discretely sampled. The frame of the estimate is based on the introduction of auxiliary data to complete the diffusion latent missing between each pair of measurements. Combine methods of Monte Carlo Markov Chain algorithm based on the Metropolis-Hastings, in conjunction with the numerical scheme of Euler-Maruyama, to sample the posterior distribution of latent data and model parameters.

Another alternative for Gaussian estimation was proposed by Nowman [2]. In this paper, an estimate using the exact discretization of the model is proposed,

$$dX_t = \alpha \left(\mu - X_t \right) dt + \sigma X_{(s\Delta)}^{\gamma} dB_t, \quad s\Delta \leq t < (s+1)\,\Delta, \tag{8}$$

assuming that volatility stays constant on the intervals $[s\Delta, (s+1)\Delta)$; $s = 0, 1, \ldots$.

In this way, we obtain an equation of the form

$$X_t = e^{-\alpha\Delta} X_{(t-\Delta)} + \mu \left(1 - e^{-\alpha\Delta} \right) + \eta_t, \tag{9}$$

where the conditional distribution $\eta_t \mid F_{t-1}^B \approx N(0, (\sigma^2/2\alpha)(1 - e^{-2\alpha\Delta})\{X_{t-1}\}^{2\gamma})$.

However, despite the obtainable approximate conditional distribution, this approach does not correspond to the Euler discretization applied to (1). This estimation technique requires the discrete model corresponding to (2) which is given by

$$X_t = e^{-\alpha} X_{(t-1)} + \mu \left(1 - e^{-\alpha} \right) + \eta_t, \quad (t = 0, 1, \ldots, T), \tag{10}$$

where η_t, $(t = 0, 1, \ldots, T)$ satisfies the conditions $E[\eta_s \eta_t] = 0$; $(s \neq t)$ and

$$E\left[\eta_t^2\right] = \int_{t-1}^{t} e^{-2(t-\tau)\alpha} \sigma^2 \{X_{(t-1)}\}^{2\gamma} d\tau$$
$$= \frac{\sigma^2}{2\alpha} \left(1 - e^{-2\alpha} \right) \{X_{(t-1)}\}^{2\gamma} = m_{tt}^2. \tag{11}$$

This states that the logarithm of the Gaussian likelihood function may be written as

$$L(\theta) = \sum_{t=1}^{T} \left[2 \log m_{tt} + \left\{ \frac{X_t - e^{-\alpha} X_{(t-1)} + \mu \left(1 - e^{-\alpha} \right)}{m_{tt}} \right\}^2 \right]$$
$$= \sum_{t=1}^{T} \left[2 \log m_{tt} + \varepsilon_t^2 \right], \tag{12}$$

where $\varepsilon = [\varepsilon_1, \ldots, \varepsilon_T]$ is a vector whose elements can be calculated.

The optimization of $L(\theta)$ is carried out using the methods developed by Bergstrom [26–30] and Nowman [2].

Yu and Phillips [1] propose a Gaussian estimation using random time changes which is based on the idea that any continuous time martingale can be written as a Brownian motion after a suitable time change. It considers the explicit solution (1) to obtain the expression

$$X_{(t+h)} = \mu \left(1 - e^{-\alpha h} \right) + e^{-\alpha h} X_{(t)} + \int_0^h \sigma e^{-\alpha(h-\tau)} \{X_{(t+\tau)}\}^{\gamma} dB_{\tau}$$
$$\text{for } h > 0. \tag{13}$$

Let $M(h) = \sigma \int_0^h e^{-\alpha(h-\tau)} \{X_{(t+\tau)}\}^{\gamma} dB_{\tau}$. $M(h)$ is a continuous martingale (In Yu and Phillips [1], $M(h)$ was assumed as a continuous martingale, but this assumption is generally not warranted, and $[M]_h$ does not induce a Brownian motion as Yu and Phillips showed in [31]. In addition, they establish that a simple decomposition splits the error process into a trend component and a continuous martingale process.) with quadratic variation

$$[M]_h = \sigma^2 \int_0^h e^{-2(h-\tau)\alpha} \{X_{(t+\tau)}\}^{2\gamma} d\tau. \tag{14}$$

They introduce a sequence of positive numbers $\{h_j\}$ which deliver the required time changes. For any fixed constant $a > 0$, let

$$h_{j+1} = \inf \left\{ s \mid \left[M_j \right]_s \geq a \right\}$$
$$= \inf \left\{ s \mid \sigma^2 \int_0^s e^{-2(s-\tau)\alpha} \left\{ X_{(t_j+\tau)} \right\}^{2\gamma} d\tau \geq a \right\}. \tag{15}$$

Then a succession of points $\{t_j\}$ is constructed using the iterations $t_{j+1} = t_j + h_{j+1}$ with $t_1 = 0$ to obtain

$$X_{(t_{j+1})} = \mu \left(1 - e^{-\alpha h_{j+1}} \right) + e^{-\alpha h_{j+1}} X_{(t_j)} + M_{h_{j+1}}, \tag{16}$$

where $M_{h_{j+1}} = B(a) \approx N(0, a)$.

Note that according to (16), (12) takes the form

$$L(\theta) = \frac{1}{N} \sum_j \left[2 \log a + \left\{ \frac{X_{(t_{j+1})} - \mu \left(1 - e^{-\alpha h_{j+1}} \right) - e^{-\alpha h_{j+1}} X_{(t_j)}}{a} \right\}^2 \right]. \tag{17}$$

Since the data are observed in discrete time, the time change formula cannot be applied directly and therefore must be approximated by

$$h_{j+1} = \Delta \min \left\{ s \mid \sum_{i=1}^{s} \sigma^2 e^{-2\alpha(s-i)\Delta} \left\{ X_{(t_j+i\Delta)} \right\}^{2\gamma} \geq a \right\}. \tag{18}$$

This should be selected so that a is the volatility of the error term unconditional in the equation

$$X_{(t+\Delta)} = \mu \left(1 - e^{-\alpha\Delta} \right) + e^{-\alpha\Delta} X_{(t)} + \varepsilon. \tag{19}$$

The implementation of this method is based on the estimate given by Nowman [2] to find the parameters estimates σ and γ. Furthermore, the estimation of the remaining parameters μ and α is carried out using the numerical optimization algorithm by Powell [32].

4. Gaussian Estimation of One-Factor Mean Reversion Processes with Constant Parameters

The development of this section will focus on the description of the proposed technique for the Gaussian estimation of mean reversion models of a single factor, with linear and nonlinear functions of diffusion constant parameters and linear trend functions with constant parameters.

As the density function for most diffusion processes is not known analytically, the proposed idea is not to find or approximate the transition density function or the marginal density function of the process. Otherwise, we use the numerical approximation of Euler on the EDE to build a new discrete process, Huzurbazar [33], and use the fact that in this approximation the normalized residuals resulting from Brownian process are independent standard normal random variables. As pointed out by Pedersen [7], Kohatsu-Higa and Ogawa [34], and Bally and Talay [35], the density function of the discrete process obtained with Euler scheme converges to the density function of the continuous process, and as discussed by Zellner [36], these residues can be interpreted as a vector of unknown parameters with the advantage of knowing their distribution. So we build a new variable for the models defined in (1) so that it can adapt to the observed data and later can be evaluated by examining the normalized residuals that will be useful in the estimation process. By knowing the density function of this new process we can find discretized likelihood function and use the conventional method of maximum likelihood estimation to find a closed formula for the parameters.

Consider the stochastic differential equation (1), with the initial condition $X_0 = x$. The terms $\alpha(\mu - X_t)$ and σX_t^γ being the trend and diffusion functions, respectively, that depend on X_t and a vector of unknown parameters $\theta = [\alpha, \mu, \sigma, \gamma]$, where $\{B_t\}_{t \geq 0}$ is a unidimensional standard Brownian motion defined on a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$. The trend function describes the movement of the process due to changes in time, while the diffusion function measures the magnitude of the random fluctuations around the trend function. Suppose that the conditions under which we guarantee the existence and uniqueness of the solution of Mao [37] are satisfied. Assume further the process observations in discrete time $X_t = X(\tau_t)$ at times $\{\tau_0, \tau_1, \ldots, \tau_T\}$ where $\Delta_t = \tau_t - \tau_{t-1} \geq 0$, $t = 1, 2, \ldots, T$. The objective is to estimate the unknown parameter vector θ given the historical information of the sample $X = (X_0, X_1, \ldots, X_T)'$ with equally spaced data. Note that the explicit solution of (1) is given by

$$X_t = \mu + (X_0 - \mu) e^{-\alpha t} + \int_0^t \sigma e^{-\alpha(t-s)} X_s^\gamma dB_s. \tag{20}$$

In the context of the likelihood, the estimation of θ is based on the likelihood function

$$\log L (X_1, X_2, \ldots, X_T \mid X_0, \theta) = \sum_{t=1}^{T} [\log g (X_{t+1} \mid X_t, \theta)], \tag{21}$$

where $g(X_{t+1} \mid X_t, \theta)$ are the Markov transition densities, or $L(\theta \mid X_1, X_2, \ldots, X_T) = f(X_1, X_2, \ldots, X_T, \theta)$ and $f(X_1, X_2, \ldots, X_T, \theta)$ is the joint density function.

Although the differential equation (1) can be solved analytically on $[0, T]$, the functions $g(X_{t+1} \mid X_t, \theta)$ and $f(X_1, X_2, \ldots, X_T, \theta)$ are not available in closed form, and, therefore, the estimate can not be accomplished, except for the case where $\gamma = 0$.

Note that the exact discretized solution of (1) takes the form

$$X (\tau_t) = \mu + (X (\tau_{t-1}) - \mu) e^{-\alpha(\tau_t - \tau_{t-1})} + \int_{\tau_{t-1}}^{\tau} \sigma e^{-\alpha(\tau_t - s)} X_{(s)}^\gamma dB_s \tag{22}$$

which can be approximated by

$$X (\tau_t) = \mu + (X (\tau_{t-1}) - \mu) e^{-\alpha \Delta_t} + \sigma e^{-\alpha \Delta_t} X^\gamma (\tau_{t-1}) \Delta B_t, \tag{23}$$

where $\Delta B_t = B_{\tau_t} - B_{\tau_{t-1}}$.

Thus,

$$X (\tau_t) = X (\tau_{t-1}) + \alpha (\mu - X (\tau_{t-1})) \Delta_t + X^\gamma (\tau_{t-1}) \varepsilon_{\tau_t}, \tag{24}$$

where $\varepsilon_{\tau_t} \sim N(0, \sigma^2 \Delta_t)$, corresponding to the numerical approximation of Euler.

Now define the new variable

$$Y_i = \frac{X_i - X_{i-1} - \alpha (\mu - X_{i-1}) \Delta}{X_{i-1}^\gamma} = \varepsilon_i, \quad \Delta \text{ constant.} \tag{25}$$

Note that Y_i, $i = 1, 2, 3, \ldots, T$ are independent random variables distributed normal, $N(0, \sigma^2 \Delta)$.

In terms of the density function can be written:

$$f (Y_1 : \theta) = \left(\frac{1}{2\pi\sigma^2 \Delta} \right)^{1/2} \times \exp \left[\frac{-1}{2\sigma^2 \Delta} \left(\frac{X_1 - X_0 - \alpha (\mu - X_0) \Delta}{X_0^\gamma} \right)^2 \right]$$

$$f(Y_2 : \theta) = \left(\frac{1}{2\pi\sigma^2\Delta}\right)^{1/2}$$

$$\times \exp\left[\frac{-1}{2\sigma^2\Delta}\left(\frac{X_2 - X_1 - \alpha(\mu - X_1)\Delta}{X_1^\gamma}\right)^2\right]$$

$$\vdots$$

$$f(Y_T : \theta) = \left(\frac{1}{2\pi\sigma^2\Delta}\right)^{1/2}$$

$$\times \exp\left[\frac{-1}{2\sigma^2\Delta}\left(\frac{X_T - X_{T-1} - \alpha(\mu - X_{T-1})\Delta}{X_{T-1}^\gamma}\right)^2\right].$$

$$(26)$$

Since each Y_i is independent can be expressed the joint density function as their marginal densities product, that is,

$$f(Y_1, Y_2, \ldots, Y_T) = f(Y_1 : \theta) f(Y_2 : \theta) \cdots f(Y_T : \theta). \quad (27)$$

Consequently, the likelihood function is given by

$$L(\theta \mid Y_T) = \left(\frac{1}{2\pi\sigma^2\Delta}\right)^{T/2}$$

$$\times \exp\left[\frac{-1}{2\sigma^2\Delta}\sum_{i=1}^{T}\left(\frac{X_i - X_{i-1} - \alpha(\mu - X_{i-1})\Delta}{X_{i-1}^\gamma}\right)^2\right]. \quad (28)$$

Taking the natural logarithm on both sides, it follows that

$$\log L = \frac{-T}{2}\log(2\pi\sigma^2\Delta)$$

$$- \frac{1}{2\sigma^2\Delta}\left[\sum_{i=1}^{T}\left(\frac{X_i - X_{i-1} - \alpha(\mu - X_{i-1})\Delta}{X_{i-1}^\gamma}\right)^2\right]. \quad (29)$$

Now consider the problem of maximizing the likelihood function given by

$$\frac{\partial \log(L)}{\partial\theta} = \vec{0}, \quad \text{i.e., } \widehat{\theta} = \underset{\theta}{\text{Arg max}}(\log(L)). \quad (30)$$

The estimation process leads to the estimated parameters α and μ, are the solutions of which the next equation system, where it is assumed that the value of γ is known. Consider

$$A - E(1 - \alpha\Delta) - \alpha\mu\Delta B = 0$$
$$C - B(1 - \alpha\Delta) - \alpha\mu\Delta D = 0, \quad (31)$$

where

$$A = \sum_{i=1}^{T}\frac{X_i X_{i-1}}{X_{i-1}^{2\gamma}}, \qquad B = \sum_{i=1}^{T}\frac{X_{i-1}}{X_{i-1}^{2\gamma}}, \qquad C = \sum_{i=1}^{T}\frac{X_i}{X_{i-1}^{2\gamma}},$$

$$D = \sum_{i=1}^{T}\frac{1}{X_{i-1}^{2\gamma}}, \qquad E = \sum_{i=1}^{T}\left(\frac{X_{i-1}}{X_{i-1}^\gamma}\right)^2. \quad (32)$$

Thus, the estimated parameters are given by

$$\widehat{\alpha} = \frac{ED - B^2 - AD + BC}{(ED - B^2)\Delta},$$

$$\widehat{\mu} = \frac{A - E(1 - \widehat{\alpha}\Delta)}{\widehat{\alpha}B\Delta}. \quad (33)$$

Additionally, the parameter σ is determined by an equation independent of the system of (31) and is completely determined by the estimated parameters:

$$\widehat{\sigma} = \sqrt{\frac{1}{T\Delta}\sum_{i=1}^{T}\left(\frac{X_i - X_{i-1} - \widehat{\alpha}(\widehat{\mu} - X_{i-1})\Delta}{X_{i-1}^\gamma}\right)^2}. \quad (34)$$

This estimation scheme has several important qualities. First, you come to a closed formula for Gaussian estimation without numerical optimization methods. Second, the computational implementation is simple and does not require sophisticated routines to perform calculations very extensive.

5. Simulations

This section will present comparisons of the results with those published by other authors using the techniques described above. Comparisons are made with the estimate based on simulated data sets to analyze statistical properties of the proposed estimators.

Yu and Phillips [1] present the results of a simulation study with the Monte Carlo technique in comparing their method with Nowman's [2]. To make this comparison are suggested three scenarios: to resolve data monthly, weekly, and daily. In all cases the basic pattern of square root is defined as

$$dX_t = (\lambda + \beta X_t)dt + \sigma X_t^\gamma dB_t;$$
$$\text{with } \lambda > 0, \ \beta < 0, \ \sigma > 0, \ \gamma = 0.5. \quad (35)$$

Note that with appropriate substitution of this model parameters are equivalent to those described in (1) (The equivalence of models (1) and (35) is satisfied by having $\lambda = \alpha\mu$ and $\beta = -\alpha$. Originally in Yu and Phillips [1], the parameter λ in (35) is defined as α but was changed to avoid confusion with the model (1).)

The specific parameters used in each stage are shown in Table 2.

To perform the first comparison parameters are set as specified in Table 2 and obtained a series simulated. Then you set the value to $\gamma = 0.5$ and proceed to the estimation by the method of Section 4 to obtain $\widehat{\lambda}$, $\widehat{\beta}$, and $\widehat{\sigma}$. This procedure was repeated 1000 times (as in Yu and Phillips [1]). Some statistical estimators are presented in Tables 3, 4, and 5.

Tables 3, 4, and 5 show that the method proposed by Nowman [2] gives good estimates for the parameters σ and γ, as manifested by Yu and Phillips [1] it becomes clear that a starting point for the method is defined by the latter. Furthermore, through the resulting estimate of the parameter σ, our method becomes much more accurate than Nowman [2], introducing a bias less than 0.2% for the monthly data series and a small value for cases of weekly and daily series.

Table 2: Monte Carlo study parameters.

	Monthly	Weekly	Daily
Size	500	1000	2000
λ	0.72	3	6
β	−0.12	−0.5	−1
σ	0.6	0.35	0.25
γ	0.5	0.5	0.5

In the case of λ and β are parameters observed inefficient performance (as discussed Yu and Phillips [1, page 220]) in the method of Nowman, λ yielding bias of 87%, 47% and 64% for data monthly, weekly and daily, respectively. Also, for β the bias present in the estimate is 94%, 46%, and 54% for the respective monthly, weekly, and daily series. The method proposed by Yu and Phillips has a better performance than Nowman, measured from the bias in the parameters, reducing it by 15%, 8%, and 16% for the case of λ and 5%, 6%, and 6% in the case of β for the respective monthly, weekly, and daily series. Our method has a performance similar to that of Yu and Phillips for daily data series, reducing the bias regarding Nowman method in 15% for λ and 4% for β. For weekly data series for reducing λ bias, regarding the method of Nowman, is 7% whereas for β is 6%, as reduction Yu and Phillips method.
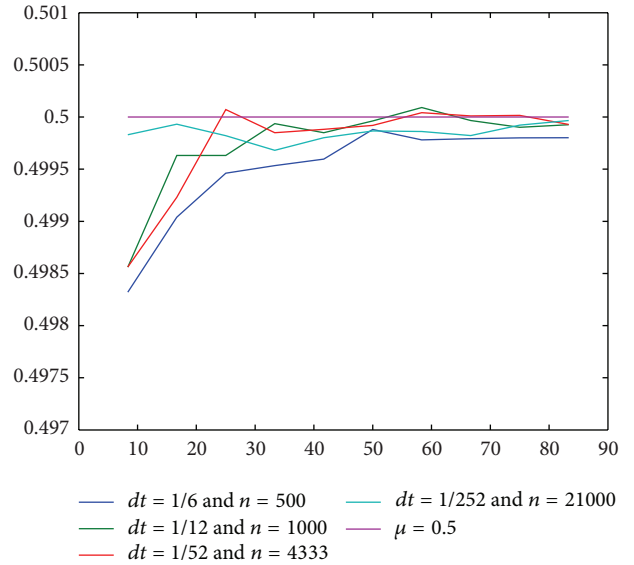
An important point to note in the results obtained by our method is the condition that the parameter γ is known a priori. This is necessary in order to find closed formulas for the estimators.

In the approach to the model proposed by Nowman and Yu-Phillips, $dX_t = (\lambda + \beta X_t)dt + \sigma X_t^\gamma dB_t$ is implied to have particular difficulty in estimating the parameter of speed of reversion. To make this condition clear, Table 6 shows the results obtained by the method of Durham and Gallant [4] and our method to estimate the parameters of the model $dX_t = \alpha(\mu + X_t)dt + \sigma X_t^\gamma dB_t$ with $\alpha > 0$, $\mu > 0$, $\sigma > 0$, $\gamma = 0.5$. The results are the average of 512 estimations of monthly series with 1000 observations.

The results shown in Table 6 show that the errors in estimation of μ and σ are small (less than 1%) for both methods. On the other hand, the estimated speed of reversion has an error of about 10% resulting in Durham and Gallant, while our method has an error close to 7%. From this analysis, it is clear that the parameter α has a higher error in estimating than the other parameter estimates.

In this regard, the estimation error of α in Yu [3] provides an analysis of the bias in the parameter estimation inherent reversion total observation time which is sampled in the continuous process, for the special case, where the long-term average μ is zero. Reference is made to the fact that the bias in parameter depends asymptotically on total observation time which of the process and not the number of observations made in this period.

Figures 1, 2, and 3 show the results of the estimation of the parameters for different observation horizons with different sampling frequencies. (The results are obtained as the average of the individual estimates on 1000 simulated
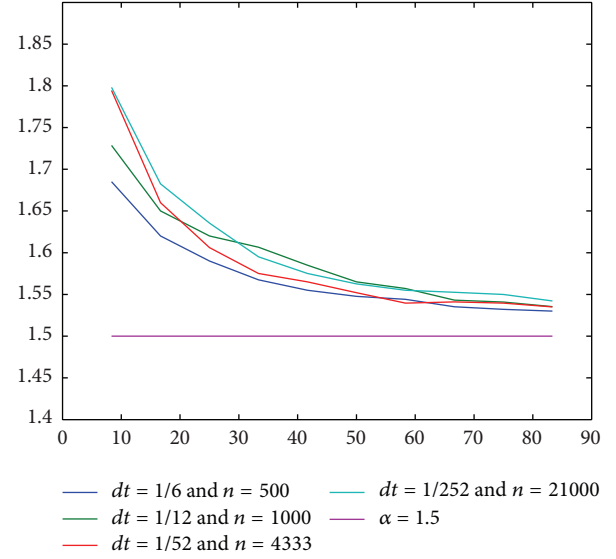
Figure 1: α estimation results. $dt = 1/6, 1/12, 1/52,$ and $1/252, n = 500, 1000, 4333,$ and 21000, respectively.

— $dt = 1/6$ and $n = 500$ — $dt = 1/252$ and $n = 21000$
— $dt = 1/12$ and $n = 1000$ — $\alpha = 1.5$
— $dt = 1/52$ and $n = 4333$

Figure 2: μ estimation results. $dt = 1/6, 1/12, 1/52,$ and $1/252, n = 500, 1000, 4333,$ and 21000, respectively.

— $dt = 1/6$ and $n = 500$ — $dt = 1/252$ and $n = 21000$
— $dt = 1/12$ and $n = 1000$ — $\mu = 0.5$
— $dt = 1/52$ and $n = 4333$

series in 4 sampling scenarios: (a) bimonthly series with 500 observations, (b) monthly series with 1000 observations; (c) weekly series with 4333 observations and (d) daily series with 21,000 observations. The reference model for the simulations was $dX_t = \alpha(\mu - X_t)dt + \sigma X_t^\gamma dB_t$, with $\alpha = 1.5$, $\mu = 0.5$, $\gamma = 0.5$, and $\sigma = 0.4$.)

Figure 1 shows the evolution of the estimate of α for observations from 10 to 83 years. Behavior of the estimates in the four scenarios can be established empirically by the dependence of the bias on the estimate and the horizon of observation as they have more years of observation

TABLE 3: Comparison of monthly data series with Nowman, Yu-Phillips, and proposed methods.

| | Nowman | | Yu, Phillips | | | | Proposed | | |
	λ	β	λ	β	σ	γ	λ	β	σ
Mean	1,3440	−0,2332	1,2330	−0,2275	0,6173	0,4919	1,2797	−0,2200	0,5989
Bias	0,6240	−0,1132	0,5130	−0,1075	0,0173	−0,0081	0,5597	−0,1000	−0,0011
Var	0,5897	0,1713	0,5732	0,1589	0,0099	0,0071	0,5406	0,0149	0,0004
MSE	0,9791	0,1841	0,8363	0,1705	0,0102	0,0071	0,8533	0,0249	0,0004

The results of the methods of Nowman and Yu, Phillips are taken into account, Yu and Phillips [1]. Our results were obtained with 1000 simulations with 500 observations each. The real values of the parameters are $\lambda = 0.72$, $\beta = −0.12$, $\sigma = 0.6$, and $\gamma = 0.5$.

TABLE 4: Comparison of weekly data series with Nowman, Yu-Phillips, and proposed methods.

| | Nowman | | Yu, Phillips | | | | Proposed | | |
	λ	β	λ	β	σ	γ	λ	β	σ
Mean	4,4090	−0,7320	4,1650	−0,7011	0,3762	0,4925	4,2038	−0,7022	0,3499
Bias	1,4090	−0,2320	1,1650	−0,2011	0,0262	−0,0075	1,2038	−0,2022	−0,0001
Var	4,0663	0,1109	3,8608	0,1030	0,0172	0,0357	3,5427	0,0992	0,0001
MSE	6,0855	0,1647	5,2180	0,1435	0,0179	0,0358	4,9883	0,1400	0,0001

The results of the methods of Nowman and Yu, Phillips are taken into account, Yu and Phillips [1]. Our results were obtained with 1000 simulations with 1000 observations each. The real values of the parameters are $\lambda = 3$, $\beta = −0.5$, $\sigma = 0.35$, and $\gamma = 0.5$.

TABLE 5: Comparison of daily data series with Nowman, Yu-Phillips, and proposed methods.

| | Nowman | | Yu, Phillips | | | | Proposed | | |
	λ	β	λ	β	σ	γ	λ	β	σ
Mean	9,8250	−1,5360	8,8440	−1,4750	0,2521	0,4970	8,9542	−1,4933	0,2500
Bias	3,8250	−0,5360	2,8440	−0,4750	0,0021	−0,0030	2,9542	−0,4933	0,0000
Var	19,1959	0,5219	17,8220	0,4798	0,0244	0,0746	15,5818	0,4334	0,0000
MSE	33,8266	0,8092	25,9100	0,7054	0,0244	0,0746	24,2938	0,6763	0,0000

The results of the methods of Nowman and Yu, Phillips are taken into account, Yu and Phillips [1]. Our results were obtained with 1000 simulations with 1000 observations each. The real values of the parameters are $\lambda = 6.0$, $\beta = −1.0$, $\sigma = 0.25$, and $\gamma = 0.5$.

TABLE 6: Comparison of Durham and Gallant, and proposed methods.

Parameter[*]	Real	Bias[a]	Bias[b]	Real	Bias[a]	Bias[b]
α	0,5	0,048900	0,037700	0,5	0,059730	0,021400
μ	0,06	0,000610	0,000127	0,06	−0,000310	−0,000021
σ	0,15	0,000200	0,000032	0,22	0,000120	−0,000046
α	0,5	0,043780	0,032400	0,4	0,045890	0,032200
μ	0,06	0,000050	0,000016	0,06	0,000770	0,000046
σ	0,03	0,000010	−0,000003	0,15	0,000080	−0,000025

The data are obtained with 512 simulations monthly series ($dt = 1/12$) with 1000 observations each, and the parameters are listed as real.
[a]Results of the method proposed by Durham and Gallant.
[b]Results of the method proposed in this paper.
[*]In Durham and Gallant [4]: $\mu = \theta_1$, $\alpha = \theta_2$, and $\sigma = \theta_3$.

bias decreases, with no evidence of a significant difference between the estimates different frequencies.

From Figure 2 it is observed that after few observations, virtually independent of the year of observation, estimations for μ are obtained with a bias of less than 1%, showing the efficiency of the estimation of the long-term average reversion processes average. In the case of estimation of σ, Figure 3 shows that the bias in the parameter estimates is more dependent on the number of observations which is the total time of observation of the process. This is reflected in the fact that the estimates are in series with a higher sampling rate reaches a lower bias for the same observation horizon than lower sample rate and therefore have fewer observations. On the other hand, as for the case of the estimate of μ, the bias present in the estimates of σ is practically negligible (less than 1%) in most scenarios.
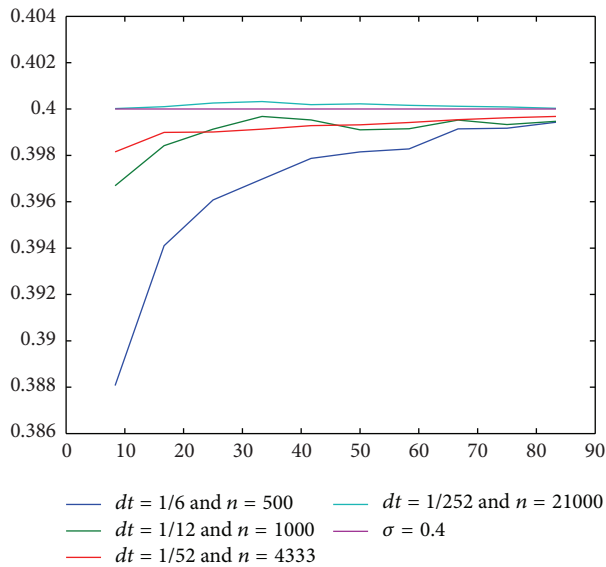
FIGURE 3: σ estimation results. $dt = 1/6$, $1/12$, $1/52$, and $1/252$, $n = 500$, 1000, 4333, and 21000, respectively.

6. Conclusions and Comments

This paper presents a new alternative for Gaussian estimation models of mean reversion of one factor under the Euler numerical approximation that includes, among others, some models of interest rates in the short term.

Monte Carlo simulations show a good approximation of the diffusion term parameter but less approximation in a trend term parameter. The estimation results have almost the same order as that in Nowman's [2] and in some cases better than that in Durham and Gallant's [4]. It should be noted that the proposed methodology outperforms the estimate of σ compared to the other alternatives presented. These results become more relevant since in practice the estimation of volatility is extremely useful in the evaluation of financial derivatives, in which the trend term loses significance when it is absorbed by the processes of change measurement and valuation risk neutral.

The estimate based on this new method, though does not eliminate the bias in estimation of the reversion rate parameter for a small sample, is in some cases comparable with the results in Yu and Phillips [1], if samples have longer time horizons.

This estimation scheme has several important qualities. First, we come to a closed formula for Gaussian estimation without numerical optimization methods. Second, the computational implementation is simpler and does not require sophisticated routines to perform extensive calculations.

References

[1] J. Yu and P. C. B. Phillips, "A Gaussian approach for continuous time models of the short-term interest rate," *The Econometrics Journal*, vol. 4, no. 2, pp. 210–224, 2001.

[2] K. Nowman, "Gaussian estimation of single-factor continuous time models of the term structure of interest rates," *The Journal of Finance*, vol. 52, pp. 1695–1703, 1997.

[3] J. Yu, "Bias in the estimation of the mean reversion parameter in continuous time models," *Journal of Econometrics*, vol. 169, no. 1, pp. 114–122, 2012.

[4] G. B. Durham and A. R. Gallant, "Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes," *Journal of Business & Economic Statistics*, vol. 20, no. 3, pp. 297–338, 2002.

[5] L. Valdivieso, W. Schoutens, and F. Tuerlinckx, "Maximum likelihood estimation in processes of Ornstein-Uhlenbeck type," *Statistical Inference for Stochastic Processes*, vol. 12, no. 1, pp. 1–19, 2009.

[6] O. Elerian, S. Chib, and N. Shephard, "Likelihood inference for discretely observed nonlinear diffusions," *Econometrica*, vol. 69, no. 4, pp. 959–993, 2001.

[7] A. R. Pedersen, "Consistency and asymptotic normality of an approximate maximum likelihood estimator for discretely observed diffusion processes," *Bernoulli*, vol. 1, no. 3, pp. 257–279, 1995.

[8] T. Koulis and A. Thavaneswaran, "Inference for interest rate models using Milstein's approximation," *The Journal of Mathematical Finance*, vol. 3, pp. 110–118, 2013.

[9] G. J. Jiang and J. L. Knight, "A nonparametric approach to the estimation of diffusion processes, with an application to a short-term interest rate model," *Econometric Theory*, vol. 13, no. 5, pp. 615–645, 1997.

[10] D. Florens-Zmirou, "On estimating the diffusion coefficient from discrete observations," *Journal of Applied Probability*, vol. 30, no. 4, pp. 790–804, 1993.

[11] O. Vasicek, "An equilibrium characterization of the term structure," *Journal of Financial Economics*, vol. 5, pp. 177–186, 1977.

[12] M. Brennan and E. Schwartz, "Analyzing convertible bonds," *Journal of Financial and Quantitative Analysis*, vol. 15, pp. 907–929, 1980.

[13] J. Cox, J. Ingersoll, and S. Ross, "An analysis of variable rate loan contracts," *The Journal of Finance*, vol. 35, pp. 389–403, 1980.

[14] K. Chan, F. Karolyi, F. Longstaff, and A. Sanders, "An empirical comparison of alternative models of short term interest rates," *The Journal of Finance*, vol. 47, pp. 1209–1227, 1992.

[15] A. Lari-Lavassani, A. Sadeghi, and A. Ware, "Mean reverting models for energy option pricing," *Electronic Publications of the International Energy Credit Association*, 2001, http://www.ieca.net.

[16] D. Pilipovic, *Energy Risk: Valuing and Managing Energy Derivatives*, McGraw-Hill, New York, NY, USA, 2007.

[17] G. Courtadon, "The Pricing of options on default-free bonds," *Journal of Financial and Quantitative Analysis*, vol. 17, pp. 75–100, 1982.

[18] J. C. Cox, J. E. Ingersoll, Jr., and S. A. Ross, "A theory of the term structure of interest rates," *Econometrica*, vol. 53, no. 2, pp. 385–407, 1985.

[19] K. Dunn and J. McConnell, "Valuation of GNMA Mortgage-backed securities," *The Journal of Finance*, vol. 36, no. 3, pp. 599–616, 1981.

[20] K. Ramaswamy and S. Sundaresan, "The valuation of floating-rate instruments," *Journal of Financial Economics*, vol. 17, pp. 251–272, 1986.

[21] S. Sundaresan, *Valuation of Swaps. Working Paper*, Columbia University, 1989.

[22] F. Longstaff, "The valuation of options on yields," *Journal of Financial Economics*, vol. 26, no. 1, pp. 97–122, 1990.

[23] G. Constantinides and J. Ingersoll, "Optimal bond trading with personal taxes," *Journal of Financial Economics*, vol. 13, pp. 299–335, 1984.

[24] G. Jiang and J. Knight, "Finite sample comparison of alternative estimators of Ito diffusion processes: a Monte Carlo study," *The Journal of Computational Finance*, vol. 2, no. 3, 1999.

[25] L. Giet and M. Lubrano, *Bayesian Inference in Reducible Stochastic Differential Equations*, Document de Travail No. 2004–57, GREQAM University d'Aix-Marseille, 2004.

[26] A. R. Bergstrom, "Gaussian estimation of structural parameters in higher order continuous time dynamic models," *Econometrica*, vol. 51, no. 1, pp. 117–152, 1983.

[27] A. Bergstrom, "Continuous time stochastic models and issues of aggregation over time," in *Handbook of Econometrics*, Z. Griliches and M. D. Intriligator, Eds., vol. 2, Elsevier, Amsterdam, The Netherlands, 1984.

[28] A. Bergstrom, "The estimation of parameters in nonstationary higher-order continuous-time dynamic models," *Econometric Theory*, vol. 1, pp. 369–385, 1985.

[29] A. Bergstrom, "The estimation of open higher-order continuous time dynamic models with mixed stock and flow data," *Econometric Theory*, vol. 2, pp. 350–373, 1986.

[30] A. Bergstrom, *Continuous Time Econometric Modelling*, Oxford University Press, Oxford, UK, 1990.

[31] P. C. B. Phillips and J. Yu, "Corrigendum to 'A Gaussian approach for continuous time models of short-term interest rates'," *The Econometrics Journal*, vol. 14, no. 1, pp. 126–129, 2011.

[32] M. J. D. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives," *The Computer Journal*, vol. 7, pp. 155–162, 1964.

[33] V. S. Huzurbazar, "The likelihood equation, consistency and the maxima of the likelihood function," *Annals of Eugenics*, vol. 14, pp. 185–200, 1948.

[34] A. Kohatsu-Higa and S. Ogawa, "Weak rate of convergence for an Euler scheme of nonlinear SDE's," *Monte Carlo Methods and Applications*, vol. 3, no. 4, pp. 327–345, 1997.

[35] V. Bally and D. Talay, "The law of the Euler scheme for stochastic differential equations: error analysis with Malliavin calculus," *Mathematics and Computers in Simulation*, vol. 38, no. 1–3, pp. 35–41, 1995.

[36] A. Zellner, "Bayesian analysis of regression error terms," *Journal of the American Statistical Association*, vol. 70, no. 349, pp. 138–144, 1975.

[37] X. Mao, *Stochastic Differential Equations and Applications*, Horwood Publishing, England, UK, 2nd edition, 2008.