*Research Article*

# Parameter Estimation in Mean Reversion Processes with Deterministic Long-Term Trend

**Freddy H. Marín Sánchez and Verónica M. Gallego**

*Department of Mathematical Sciences, Universidad Eafit, Carrera 49 No. 7 Sur 50, Medellin, Colombia*

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

This paper describes a procedure based on maximum likelihood technique in two phases for estimating the parameters in mean reversion processes when the long-term trend is defined by a continued deterministic function. Closed formulas for the estimators that depend on observations of discrete paths and an estimation of the expected value of the process are obtained in the first phase. In the second phase, a reestimation scheme is proposed when *a priori* knowledge exists of the long-term trend. Some experimental results using simulated data sets are graphically illustrated.

## 1. Introduction

In the literature it is common to find studies and applications of mean reversion models, at both practical and theoretical levels, especially in fields such as energy markets, commodities, and interest rates. In these models there is a long-term trend which acts as an attractor making the process to oscillate around it and there is a random component that adds volatility to the movement. The specific characteristics of the models will depend on the structure of the trend and the volatility of the process.

Some common models of mean reversion are those in which all the parameters in the process are constant, such as Ornstein-Uhlenbeck model, the CIR model proposed by Cox et al. [1], and in general the models obtained from the CKLS structure proposed by Chan et al. [2]. In addition to these models there are other methods not so common in the literature in which the reversion is determined not by a parameter but a deterministic function or another stochastic process [3, 4]. In general, the fit of the models to specific situations is done by assigning values to the parameters and functions, either from *a priori* knowledge of the problem or from parameter estimation techniques based on historical

data series that help to detect unobservable information from the process.

One of the first difficulties in the application of parameter estimation techniques is that it is necessary to have information with the same time resolution that is specified in the model. Since models of stochastic differential equations are formulated in continuous time and that the observed paths of the process can be obtained only in discrete time, it is necessary to discretize the EDE model. An approach frequently used for this purpose is given by Euler-Maruyama. With this methodology a new discrete-time process is obtained, with which it is possible to make inferences that are valid for continuous time model, as in the case of estimating the parameters.

There are different procedures that can be implemented such as methods based on distributional moments [5], Kernel methods [6, 7], least squares, and maximum likelihood [8–10]. Parameter estimation becomes fundamental too for modeling control systems. In the literature it is possible to find different methodologies for those issues such as those developed for Meng and Ding [11], where the parameter estimation in an equation-error autoregressive (EEAR) system is done through a transformation of the model to an

equivalent one removing the autoregressive term and to obtain the parameters of the original model the principle of equivalence is used. Another methodology is proposed for Cao and Liu [12] where the parameter estimation in a power system is carried out from a recursive procedure to estimate simultaneously all the parameters from techniques based on the hierarchical identification principle, using gradient algorithms and least squared algorithms. Complementarily, Ding et al. [13] used methods based on the gradient algorithms and least squares iterative algorithms for system identification in output error (OE) and output error moving average (OEMA) systems, where the parameters that depend on unknown variables are computed using estimates of these unknown variables from previously estimated parameters.

In the area of multirate system identification, Ding et al. [14, 15] used the polynomial transformation technique to deal with the identification problem for dual-rate stochastic systems, while Sahebsara et al. [16] discussed the parameter estimation of multirate multi-input multioutput systems. Also, Ding and Chen [17] presented the combined parameter and state estimation algorithms of the lifted state-space models for general dual-rate systems based on the hierarchical identification method. Shi et al. [18] gave a crosstalk identification algorithm for multirate xDSL FIR systems.

In general, the auxiliary model identification idea has been used to solve the identification problem of dual-rate single-input single-output systems as shown in [19].

For more detailed information about multi-innovation stochastic gradient algorithm for multi-input multioutput systems and for multirate multi-input systems as well as auxiliary models based multi-innovation extended stochastic identification theory (see [20] and references therein). We will concentrate on one-factor stochastic model in which the parameters are estimated using maximum likelihood based on the discretized model when the long-term trend is given by a deterministic function. In this case we must estimate both the trend function and the parameters. To estimate the long-term trend function some convolution techniques and numerical differentiation are used, whereas the normality properties of residuals resulting from discretization are used for estimating parameters by maximum likelihood and in this way it is possible to obtain closed formulas for estimators based on the observations of a path of the process and the estimation of the long-term trend.

In the estimation process there may be some bias in the estimates with respect to the actual values despite the fact that they are obtained by the method of maximum likelihood. When this situation occurs it is necessary to develop alternative methodologies to have a greater adjustment to the estimates based on the initial estimates [21].

This paper is organized as follows. In Section 2 a description of mean reversion processes is presented for cases when the long-term trend is given by a continued deterministic function. The first phase of the estimation technique is shown in Section 3. Section 4 presents some examples and preliminary results for the estimation technique and in Section 5 (second phase), a procedure for reestimating the parameters from additional *a priori* knowledge is described.

## 2. Mean Reversion Processes with Deterministic Long-Term Trend

Mean reversion processes of one factor with constant parameters and mean given by a deterministic function can be written as

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

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

$\alpha$ parameter is called reversion rate, $\mu(t)$ is the mean reversion level or trend of long-run equilibrium, $\sigma$ is the parameter associated with the volatility, and $\gamma$ determines the sensitivity of the variance to the level of $X_t$.

Model (1) is a generalization of the models CKLS, Chan et al. [2], where the mean reversion level is a deterministic function that captures the trend of the process. In a general way, $\mu(t)$ plays the role of an attractor at each point $t$ in the sense that, when $X_t > \mu(t)$ the trend term $\alpha(\mu(t) - X_t) < 0$ and therefore $X_t$ decreases and when $X_t < \mu(t)$ a similar argument establishes that $X_t$ grows.

To establish the relation between the mean reversion function $\mu(t)$ and the expected value of the process $E[X_t]$, we can consider (1) written in integral form as

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

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

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

The solution of this equation is $m(t) = m(0)e^{-\alpha t} + e^{-\alpha t} \int_0^t e^{\alpha s} \mu(s) ds$.

To illustrate graphically the relation between $\mu(t)$ and $m(t)$, we use $\mu(t) = a + b \sin(2\pi t/P)$ as an example. In this case the solution to (3) is given by

$$m \left( t \right) = a + \left[ m_0 - a + \frac{\alpha \beta b}{\alpha^2 + \beta^2} \right] e^{-\alpha t}$$

$$+ \frac{\alpha b}{\alpha^2 + \beta^2} \left[ \alpha \sin \left( \beta t \right) - \beta \cos \left( \beta t \right) \right], \qquad (4)$$

where $\beta = 2\pi/P$.

Figure 1 shows the dynamic behavior of $\mu(t)$, $m(t)$ and one path of $X_t$ for some specific values in the parameters.

From the Figure it can be seen that the expected value of the process mimics the long-term mean: that is just the way mean reversion works.

## 3. Phase 1: Parameter Estimation

Similar to the procedure established in Marín Sánchez and Palacio [10] and Huzurbazar [22] the main objective is to obtain closed formulas for the parameter estimators from
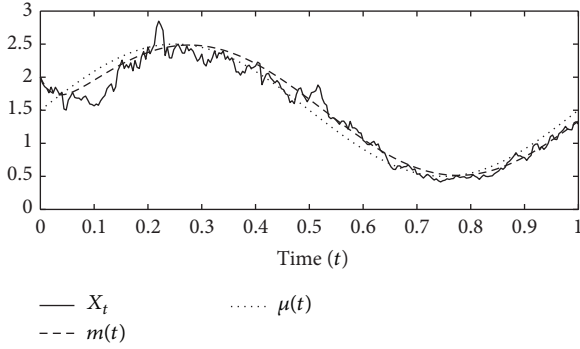
FIGURE 1: Dynamic behavior for $m(t)$, $\mu(t)$, and $X_t$, where $X_0 = 2$, $\alpha = 37$, $\sigma = 0.7$, $\gamma = 1$, $a = 1.5$, $b = 1$, $P = 1$, and $\beta = 2\pi$ with $\Delta t = 1/250$ for a total of 250 observations.

discrete observations of a sample path of the process. This can be achieved from the discretized maximum likelihood function, which is obtained from the process that results from the Euler-Maruyama discretization scheme applied to (1) and using the normal and independent increments properties of the Brownian Motion over the residuals of the discrete process.

Consider the differential equation (1) with the initial value $X(0) = x$. $\alpha(\mu(t) - X_t)$ and $\sigma X_t^\gamma$ are, respectively, the trend and diffusion functions that depend on $t$, $X_t$, and a vector of unknown parameters $\theta = [\alpha, \sigma, \gamma]$. Now suppose that the conditions that guarantee the existence and uniqueness of the solution Mao [23] are satisfied, so the expected value of $X_t$ exists and $\mu(t)$ can be defined using $m(t)$ and its derivative $\dot{m}(t)$ (see (3)). In this way,

$$\mu(t) = m(t) + \frac{\dot{m}(t)}{\alpha}. \tag{5}$$

Replacing $\mu(t)$ in (1) results in

$$dX_t = \alpha\left(m(t) + \frac{\dot{m}(t)}{\alpha} - X_t\right)dt + \sigma X_t^\gamma dB_t;$$

$$t \in [0, T]. \tag{6}$$

Now we can assume that we have the discrete observations $X_t = X(\tau_t)$ for the times $\{\tau_0, \tau_1, \ldots, \tau_T\}$, where $\Delta_t = \tau_t - \tau_{t-1} \geq 0$, $t = 1, 2, \ldots, T$. Using the numerical scheme of Euler-Maruyama on (6) it follows that

$$X(\tau_t) = X(\tau_{t-1})$$
$$+ \left[\alpha\left(m(\tau_{t-1}) - X(\tau_{t-1})\right) + \dot{m}(\tau_{t-1})\right]\Delta_t \tag{7}$$
$$+ X^\gamma(\tau_{t-1})\varepsilon_{\tau_t},$$

where $\varepsilon_{\tau_t} \approx N(0, \sigma^2\Delta_t)$.

Define the new variable $Y_t$:

$$Y_t = \frac{X_t - X_{t-1} - \left[\alpha\left(m_{t-1} - X_{t-1}\right) + \dot{m}_{t-1}\right]\Delta}{X_{t-1}^\gamma} \tag{8}$$

with $\Delta$ constant.

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

The variable $Y_t$ depends on the observations $X = (X_0, X_1, \ldots, X_T)'$ of the sample path of the process and on the unobservable quantities $m(t)$ and its derivative $\dot{m}(t)$. Taking into account that the sample path resumes all the information that is known from the process, it is necessary to estimate the expected value $m(t)$ from those observations. As in Tifenbach [4], p. 31-32, we can assume that the expected value of the process can be approximated using a convolution[1] over the sample path, so the variable $m(t)$ will depend only on the observations $X = (X_0, X_1, \ldots, X_T)'$.

Thus, we assume that the expected value of the process $X_t$ can be approximated by taking a convolution of the sample path $X = (X_0, X_1, \ldots, X_T)$ for some particular convolution function $c(t)$ given by

$$m_i = \sum_{k=-N}^{N} c_k X_{i-k}. \tag{9}$$

The most common convolution is the moving average, where $c_k$ constants are given by $c_k = 1/(2N + 1)$ for all $N$, which is denoted by $(c_t^1)$. Other convolutions can be defined varying the relative weight of each observation: for example, they can be defined from the coefficients of odd rows of Pascal's triangle which is denoted by $(c_t^2)$. In this case the central observations have a greater weight than more distant observations and for this reason the convolution is close to the path. Those specific cases of convolutions are presented in Figure 2.

The convolution is used for the purpose of smoothing the sample path, eliminating the noise and capturing the trend. This objective can also be achieved using other techniques like filtering. In Figure 2 the Hodrick-Prescott [24] filter denoted by $(c_t^{hp})$ was used to obtain an estimation of $m(t)$ that are smoother than the convolutions $(c_t^1)$ and $(c_t^2)$.

At this point we have that the expected value $m(t)$ can be approximated from the sample path $X = (X_0, X_1, \ldots, X_T)$. Now, to get the derivative of $m(t)$ we can apply numerical derivatives techniques like Taylor's theorem: for example, the derivative at the observation $i$ can be defined using a three-point rule as

$$\dot{m}_i = \frac{2m_{i+1} - 3m_i + m_{i-1}}{\Delta}, \tag{10}$$

for $i = 1, 2, \ldots, T - 1$. For $i = 0$ the derivate is given by $\dot{m}_0 = (m_1 - m_0)/\Delta$ and for $i = T$ by $\dot{m}_T = (m_T - m_{T-1})/\Delta$. This procedure can be done using a different number of points to calculate the derivate at each point.

With the estimation of $m(t)$ and $\dot{m}(t)$ we have one realization of $Y_t$ and now we can define its normal density function:

$$f(Y_i : \theta) = \left(\frac{1}{2\pi\sigma^2\Delta}\right)^{1/2}$$
$$\cdot \exp\left[\frac{-1}{2\sigma^2\Delta}\left(\frac{X_i - X_{i-1} - \left[\alpha\left(m_{i-1} - X_{i-1}\right) + \dot{m}_{i-1}\right]\Delta}{X_{i-1}^\gamma}\right)^2\right]. \tag{11}$$
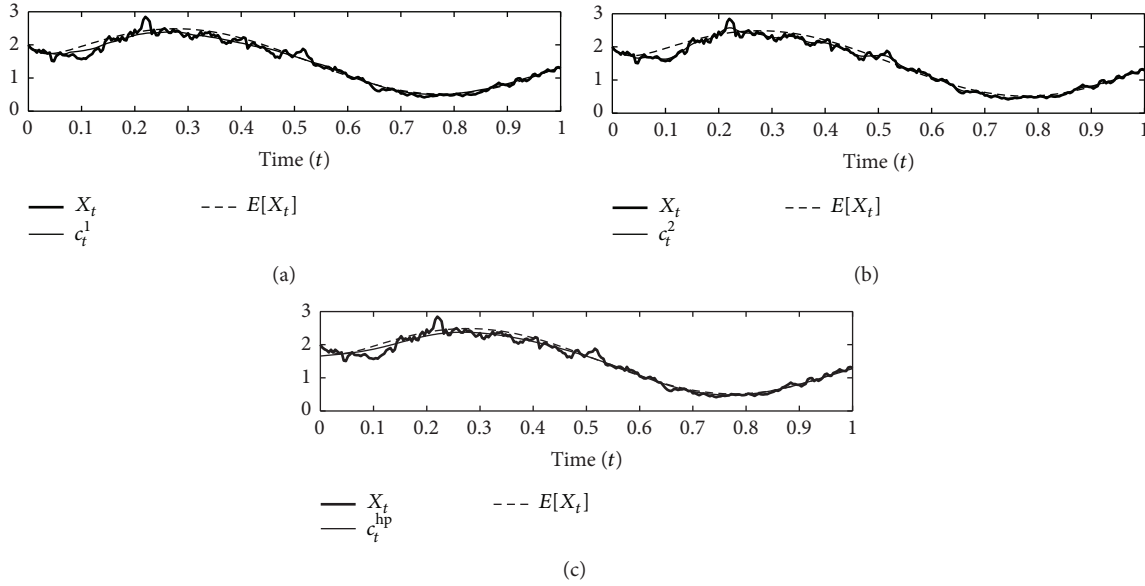
FIGURE 2: Dynamic behavior for process $X_t$, the expected value $m(t)$, and three types of convolution. (a) Moving average ($c_t^1$) with $N = 20$. (b) Pascal's weights ($c_t^2$) with $N = 20$. (c) Hodrick-Prescott filter ($c_t^{hp}$) with $\lambda_{hp} = 40000$. For all the cases $\Delta t = 1/250$ for a total of 250 observations and the parameters of $X_t$ are the same as in Figure 1.

Since each $Y_i$ is independent the joint density function can be expressed as the product of their marginal densities:

$$f\left(Y_1, Y_2, \ldots, Y_T\right) = \prod_{k=1}^{T} f\left(Y_k : \theta\right). \tag{12}$$

Consequently, the likelihood function is given by

$$L\left(\theta \mid \{Y_t\}\right) = \left(\frac{1}{2\pi\sigma^2\Delta}\right)^{T/2} \exp\left[\frac{-1}{2\sigma^2\Delta}\right.$$

$$\left. \cdot \sum_{i=1}^{T}\left(\frac{X_i - X_{i-1} - [\alpha\left(m_{i-1} - X_{i-1}\right) + \dot{m}_{i-1}]\Delta}{X_{i-1}^\gamma}\right)^2\right]. \tag{13}$$

Taking the natural logarithm on both sides it follows that

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

$$- \frac{1}{2\sigma^2\Delta}\left[\sum_{i=1}^{T}\left(\frac{X_i - X_{i-1} - [\alpha\left(m_{i-1} - X_{i-1}\right) + \dot{m}_{i-1}]\Delta}{X_{i-1}^\gamma}\right)^2\right]. \tag{14}$$

Now consider the problem of maximizing the likelihood function given by

$$\frac{\partial \log(L)}{\partial \theta} = \vec{0}, \quad \text{ie}, \widehat{\theta} = \arg\max_{\theta}\left(\log(L)\right). \tag{15}$$

If we assumed that the value of $\gamma$ is known *a priori*, the estimation process leads to the estimated parameter $\alpha$ given by

$$\widehat{\alpha} = \frac{\sum_{i=1}^{T}\left(\left(X_i - X_{i-1} - \dot{m}_{i-1}\Delta\right)\left(m_{i-1} - X_{i-1}\right)/X_{i-1}^{2\gamma}\right)}{\sum_{i=1}^{T}\left[\left(m_{i-1} - X_{i-1}\right)/X_{i-1}^\gamma\right]^2\Delta}. \tag{16}$$

The estimator $\widehat{\sigma}$ is determined by the equation:

$$\widehat{\sigma}$$

$$= \sqrt{\frac{1}{T\Delta}\sum_{i=1}^{T}\left(\frac{X_i - X_{i-1} - [\widehat{\alpha}\left(m_{i-1} - X_{i-1}\right) + \dot{m}_{i-1}]\Delta}{X_{i-1}^\gamma}\right)^2}. \tag{17}$$

Finally, with the estimation of $\widehat{\alpha}$ we can now approximate $\widehat{\mu}(t)$ with (5).

## 4. Preliminary Results

The aim of this section is to show the performance of the estimators defined in the previous section. To make this we will simulate some paths of a process satisfying (1) with a set of parameters $\alpha$ and $\sigma$ and giving a deterministic form to the $\mu(t)$ function and then find the estimators with the established procedure in order to determine—using statistical properties—if the estimates are approximate to actual parameters.

We analyze the performance of the estimators for sinusoidal and parabolic $\mu(t)$ trend functions and for a set of values of $\alpha$ and $\sigma$ when $\gamma = 1$. Each generated path consists of 250 observations with $\Delta t = 1/250$. The statistics are calculated for 1000 different simulations. Considering that for the parameters estimation it is necessary to find the expected value of the process from each path, we compare two different convolution functions and the Hodrick-Prescott filter. For numerical derivation a five-point derivation rule is used.

The values of the parameters and $\mu(t)$ functions used in the simulations are presented in Table 1.

In Table 2 we present some basic statistics (mean $\overline{\theta}$ and standard deviation $s_\theta$) for the estimators obtained from

Table 1: Simulation parameters for sinusoidal and parabolic trend functions.

|  | Case 1 | Case 2 |
| --- | --- | --- |
| $\alpha$ | 30 | 30 |
| $\sigma$ | 0.4 | 0.13 |
| $X_0$ | 2 | 1.5 |
| $\mu(t)$ | $1.5 + \sin(2\pi t)$ | $-0.5t^2 + t + 1.5$ |

the simulations. The convolution functions used in the simulations are the ones defined for the weights $(c_k^1)$ and $(c_k^3)$. $c_k^1$ weights are the same defined in Section 3 for the moving average and $c_k^3$ weights are given by $c_{k,N}^3 = (2N - |k|)/N(3N + 1)$ for $k$ from $-N$ to $N$. In this convolution the central observation has a relative weight twice the most distant observation.

From Table 2 we can conclude that the estimation technique is working well, especially in the first example (sinusoidal trend). The relative error of the estimators is less than 2% in the case of the parameter $\sigma$ but there is a bias in the parameter $\alpha$ that depends on the convolution function and at the same time on the parameter $N$ used in such convolution.

In Figure 3 we can see the performance of the estimation of the deterministic function $\mu(t)$. In that figure an estimation from a random path is shown instead of the average. The estimates are close to the theoretical function in both cases but having some differences at the beginning and the final of the observations. One form to measure the fit between the theoretical function and its estimation is using the Root Mean Square error (RMS). Table 3 presents that measure for the proposed estimation in cases 1 and 2.

## 5. Phase 2: Re-Estimation

From the simulations we can see that the estimate of $\mu(t)$ is not very accurate and that differences come from $\widehat{m}(t)$ and $\widehat{\dot{m}}(t)$, and in turn this affects the estimation of other parameters.

If it were possible to have all the paths of the process it would not be necessary for numerical approximation of the expected value or its derivative to capture information from $\mu(t)$, and it could be enough to perform an average of all paths at each time $t_i$ and that could generate pretty accurate estimations. As this condition is not the usual but only has a path, the procedure detailed above is necessary and this involves the errors mentioned in the estimates of $m(t)$ and $\dot{m}(t)$.

To correct this deficiency it is proposed to carry out a reestimate procedure trying to adjust a function to the initial estimate from *a priori* knowledge of the process. This recalculation allows a correction in the other estimators while providing a functional form that allows simulations of future paths of the process.

The following procedure is proposed to perform the reestimate of $\mu(t)$:

(1) Get an estimate of $\widehat{\mu}(t)$ according to the procedure described in Section 4.

(2) Define a functional structure $f(\Theta, t)$ from *a priori* knowledge of the process.

(3) Estimate the vector parameters $\Theta$ minimizing the error function defined by

$$\text{Error}(\Theta) = \sum_{t=1}^{T} \left[ f(\Theta, t) - \widehat{\mu}(t) \right]^2. \quad (18)$$

As an example, suppose the functional form of the long-term mean is defined as

$$f(\Theta, t) = at^2 + bt + c, \quad (19)$$

where $\Theta = [a, b, c]'$. In this case the error function would be given by

$$\text{Error}(\Theta) = \sum_{t=0}^{T} \left[ \widehat{\mu}(t) - \left( at^2 + bt + c \right) \right]^2. \quad (20)$$

To estimate $\Theta$ it is necessary to solve the system of equations defined by $\nabla \text{Error}(\Theta) = 0$, which is equivalent to the system of linear equations $M\Theta = d$, where

$$M = \begin{bmatrix} \sum t^4 & \sum t^3 & \sum t^2 \\ \sum t^3 & \sum t^2 & \sum t \\ \sum t^2 & \sum t & (T+1) \end{bmatrix}$$

$$d = \begin{bmatrix} \sum t^2 \widehat{\mu}(t) \\ \sum t \widehat{\mu}(t) \\ \sum \widehat{\mu}(t) \end{bmatrix}. \quad (21)$$

A similar procedure can be implemented under the assumption that $f(\Theta, t) = a + b\sin((2\pi/P)t)$, to find $\Theta = [a, b]'$ if $P$ is known. A comparison between $\mu(t)$, $\mu_{\text{est}}(t)$, and $\mu_{\text{rest}}(t) = f(\widehat{\Theta}, t)$ is shown in Figure 4 for both cases and for a random path of the process defined with the same set of parameters of Table 1.

Having a best estimate of $\mu(t)$ allows a reestimation of the parameters $\alpha$ and $\sigma$. To achieve this we must perform a similar procedure to that proposed in Section 4, defining

$$Y_t = \frac{X_t - X_{t-1} - \alpha(\mu_{t-1} - X_{t-1})\Delta}{X_{t-1}^{\gamma}} \quad \Delta \text{ constant.} \quad (22)$$

So,

$$\widehat{\alpha} = \frac{\sum_{i=1}^{T} \left( (X_i - X_{i-1})(\widehat{\mu}_{i-1} - X_{i-1})\Delta / X_{i-1}^{2\gamma} \right)}{\sum_{i=1}^{T} \left[ (\widehat{\mu}_{i-1} - X_{i-1}) / X_{i-1}^{\gamma} \right]^2 \Delta},$$

$$\widehat{\sigma} = \sqrt{\frac{1}{T\Delta} \sum_{i=1}^{T} \left( \frac{X_i - X_{i-1} - \widehat{\alpha}(\widehat{\mu}_{i-1} - X_{i-1})\Delta}{X_{i-1}^{\gamma}} \right)^2}. \quad (23)$$

Table 2: Results obtained with 1000 simulations with 250 observations each. The real values of the parameters are shown in Table 1. For case 1 in the convolutions we use $N = 22$ and for case 2 $N = 60$. For both cases we set $\lambda = 62500$ in $c_k^{hp}$.

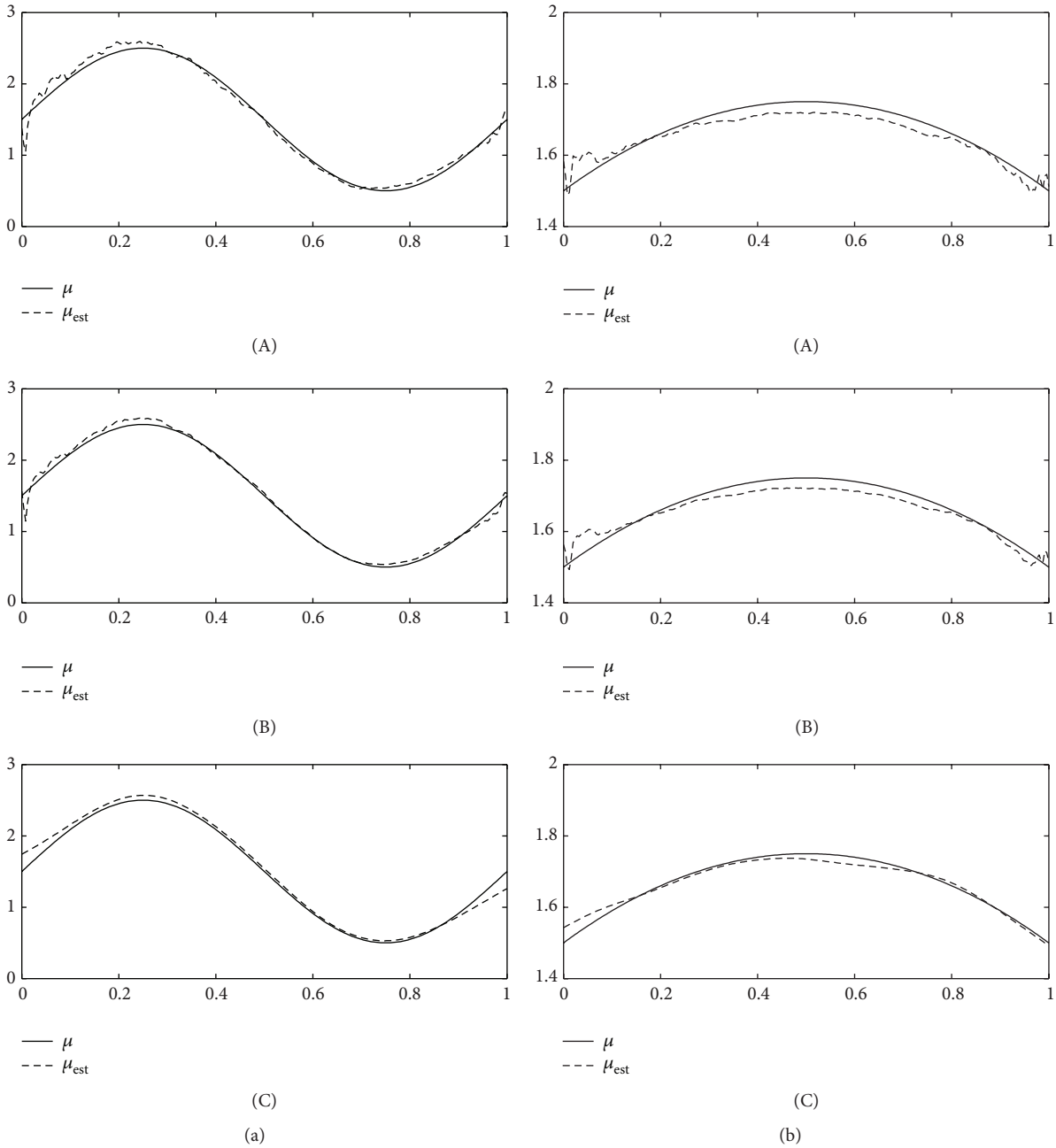| | | Case 1 | | | | | Case 2 | | |
|---|---|---|---|---|---|---|---|---|---|
| $\mu_1(t)$ | | $c_k^1$ | $c_k^3$ | $c_k^{hp}$ | $\mu_2(t)$ | | $c_k^1$ | $c_k^3$ | $c_k^{hp}$ |
| $\alpha$ | $\overline{\alpha}$ | 24.8540 | 31.3347 | 32.8771 | $\alpha$ | $\overline{\alpha}$ | 31.1160 | 35.8249 | 51.1093 |
| | $s_\alpha$ | 3.2218 | 3.9947 | 5.1406 | | $s_\alpha$ | 5.8254 | 6.6984 | 10.1414 |
| $\sigma$ | $\overline{\sigma}$ | 0.4002 | 0.3942 | 0.4023 | $\sigma$ | $\overline{\sigma}$ | 0.1293 | 0.1281 | 0.1267 |
| | $s_\sigma$ | 0.0177 | 0.0174 | 0.0174 | | $s_\sigma$ | 0.0060 | 0.0059 | 0.0060 |



Figure 3: Estimated $\mu_1(t)$ and $\mu_2(t)$. The left side correspond to Case 1 (sinusoidal) and the right side to Case 2 (parabolic). The (A) plots are obtained with the convolution ($c_k^1$), (B) plots with ($c_k^3$) and (C) plots with ($c_k^{hp}$). The parameters involved in this figure are presented in Tables 1 and 2.
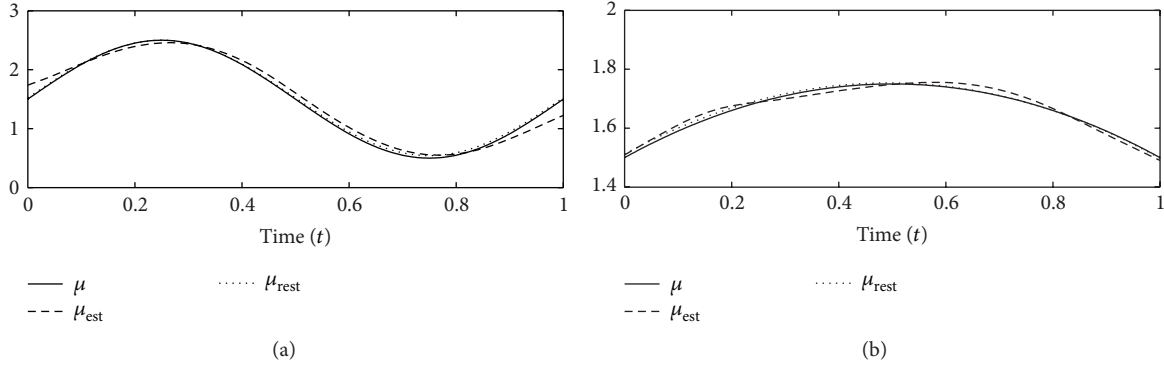
FIGURE 4: Comparison of dynamic behavior of $\mu(t)$, $\mu_{\text{est}}(t)$, and $\mu_{\text{rest}}(t)$. In case (a) $\mu_1(t) = 1.5 + \sin(2\pi t)$ and in case (b) $\mu_2(t) = -0.5t^2 + t + 1.5$.

TABLE 3: Results of RMS error for each estimation $\mu_1(t)$ and $\mu_2(t)$.

| RMS | Case 1 | Case 2 |
| --- | --- | --- |
| $c_k^1$ | 0.0782 | 0.0266 |
| $c_k^3$ | 0.0578 | 0.0238 |
| $c_k^{\text{hp}}$ | 0.0767 | 0.0140 |

Table 4 summarizes the results obtained from the first estimation and the reestimation procedure for the cases where $\mu(t)$ has a sinusoidal form and a parabolic form and the other parameters are defined in Table 1.

From the results of the above example we can observe how the reestimation procedure reduces the bias in the initial estimation of $\alpha$ for both sinusoidal and parabolic cases. The final bias in $\alpha$ is approximately 1.5% of the real value of the parameter, reducing almost 90% of the error in the first estimation in the case of sinusoidal function and near 60% for the parabolic function. The reestimations of $\sigma$ do not significantly change from the original estimate. For the case of $\mu(t)$ the RMS error is present in Table 5 for a random path of the process. In both cases the error in the reestimation is less than the 50% of the error present in the initial estimation.

## 6. Conclusions and Comments

This paper shows how starting from a path of a stochastic process is possible to find an approximation of the parameters that define the underlying dynamics It is important to highlight that, for modeling the dynamic behavior of prices of some commodities or interest rates, only one realization of the process in discrete time is available and it represents the only source of information to capture the nature of the process. Therefore, it becomes very important to obtain and interpret the invisible and visible information found on the path to make projections more adjusted to the future process behavior being studied.

It is essential that a model that seeks to reproduce the behavior of a process that is generated in the real world incorporates *a priori* information provided by an expert in the topic and that this additional information allows the model to better fit the process. This prior information is key in the development of the methodology proposed in this paper,

TABLE 4: Estimated and reestimated parameters. The estimations are calculated with the convolution ($c_k^{\text{hp}}$). Results obtained with 1000 simulations with 250 observations each for $\mu_1$ and $\mu_2$ as defined in Table 1: case 1 $\alpha = 30$ and $\sigma = 0.4$, case 2 $\alpha = 30$ and $\sigma = 0.13$.

| $\mu_1(t)$ | | $\widehat{\alpha}$ | $\widehat{\widehat{\alpha}}$ | $\mu_2(t)$ | | $\widehat{\alpha}$ | $\widehat{\widehat{\alpha}}$ |
| --- | --- | --- | --- | --- | --- | --- | --- |
| $\alpha$ | $\overline{\alpha}$ | 24.8540 | 29.6620 | $\alpha$ | $\overline{\alpha}$ | 31.1160 | 30.4616 |
| | $s_\alpha$ | 3.2218 | 3.4518 | | $s_\alpha$ | 5.8254 | 5.6828 |
| $\mu_1(t)$ | | $\widehat{\sigma}$ | $\widehat{\widehat{\sigma}}$ | $\mu_2(t)$ | | $\widehat{\sigma}$ | $\widehat{\widehat{\sigma}}$ |
| $\sigma$ | $\overline{\sigma}$ | 0.4002 | 0.4021 | $\sigma$ | $\overline{\sigma}$ | 0.1293 | 0.1300 |
| | $s_\sigma$ | 0.0177 | 0.0187 | | $s_\sigma$ | 0.0060 | 0.0060 |

TABLE 5: Results of RMS error for the estimation and reestimation of $\mu(t)$: case 1 sinusoidal trend and case 2 parabolic trend.

| RMS | Case 1 | Case 2 |
| --- | --- | --- |
| $\mu_{\text{est}}(t)$ | 0.0961 | 0.0137 |
| $\mu_{\text{rest}}(t)$ | 0.0261 | 0.0058 |

because it allows us to adjust the initial estimates of the parameters making more reliable the results obtained with the adjusted models.

On the other hand, in the development of the proposed methodology it can be seen that it allows us to obtain closed formulas for the maximum likelihood estimators in mean reversion process when the long-term trend is defined by a deterministic function, decreasing computational effort and allowing greater understanding of procedures.

## Competing Interests

The authors declare that they have no competing interests.

## Endnotes

1. A convolution function $c(t)$ is an even, continuous, and real valued function satisfying the fact that $c(t) \geqslant 0$ and $\int_{-\infty}^{\infty} c(t)dt = 1$. A convolution of a curve $s(t)$ is another function $l(t)$ defined by

$$l(t) = \int_{-\infty}^{\infty} c(w)s(t-w)\,dw, \qquad (*)$$

where $c(t)$ is a convolution function.

In the discrete case the convolution of two sequences can be written as

$$l_i = \sum_{k=-\infty}^{\infty} c_k s_{i-k}. \qquad (**)$$

## References

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

[2] K. C. Chan, G. A. Karolyi, F. A. Longstaff, and A. B. Sanders, "An empirical comparison of alternative models of the short-term interest rate," *The Journal of Finance*, vol. 47, no. 3, pp. 1209–1227, 1992.

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

[4] B. Tifenbach, *Numerical Methods for Modeling Energy Spot Prices*, University of Calgary, 2000.

[5] T. Koulis and A. Thavaneswaran, "Inference for interest rate models using Milstein's approximation," *Journal of Mathematical Finance*, vol. 3, no. 1, pp. 110–118, 2013.

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

[7] 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.

[8] K. B. Nowman, "Gaussian estimation of single-factor continuous time models of the term structure of interest rates," *The Journal of Finance*, vol. 52, no. 4, pp. 1695–1706, 1997.

[9] J. Yu and P. Phillips, "Corrigendum to 'a Gaussian approach for continuous time models of the shortterm interest rate'," *Econometrics Journal*, vol. 14, pp. 126–129, 2011.

[10] F. H. Marín Sánchez and J. S. Palacio, "Gaussian estimation of one-factor mean reversion processes," *Journal of Probability and Statistics*, vol. 2013, Article ID 239384, 10 pages, 2013.

[11] D. Meng and F. Ding, "Model equivalence-based identification algorithm for equation-error systems with colored noise," *Algorithms*, vol. 8, no. 2, pp. 280–291, 2015.

[12] Y. Cao and Z. Liu, "Signal frequency and parameter estimation for power systems using the hierarchical identification principle," *Mathematical and Computer Modelling*, vol. 52, no. 5-6, pp. 854–861, 2010.

[13] F. Ding, P. X. Liu, and G. Liu, "Gradient based and least-squares based iterative identification methods for OE and OEMA systems," *Digital Signal Processing*, vol. 20, no. 3, pp. 664–677, 2010.

[14] F. Ding, P. X. Liu, and Y. Shi, "Convergence analysis of estimation algorithms for dual-rate stochastic systems," *Applied Mathematics and Computation*, vol. 176, no. 1, pp. 245–261, 2006.

[15] F. Ding, P. X. Liu, and H. Z. Yang, "Parameter identification and intersample output estimation for dual-rate systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, vol. 38, no. 4, pp. 966–975, 2008.

[16] M. Sahebsara, T. Chen, and S. L. Shah, "Frequency-domain parameter estimation of general multi-rate systems," *Computers and Chemical Engineering*, vol. 30, no. 5, pp. 838–849, 2006.

[17] F. Ding and T. Chen, "Hierarchical identification of lifted state-space models for general dual-rate systems," *IEEE Transactions on Circuits and Systems. I. Regular Papers*, vol. 52, no. 6, pp. 1179–1187, 2005.

[18] Y. Shi, F. Ding, and T. Chen, "Multirate crosstalk identification in xDSL systems," *IEEE Transactions on Communications*, vol. 54, no. 10, pp. 1878–1886, 2006.

[19] F. Ding and T. Chen, "Combined parameter and output estimation of dual-rate systems using an auxiliary model," *Automatica*, vol. 40, no. 10, pp. 1739–1748, 2004.

[20] L. Han, J. Sheng, F. Ding, and Y. Shi, "Auxiliary model identification method for multirate multi-input systems based on least squares," *Mathematical and Computer Modelling*, vol. 50, no. 7-8, pp. 1100–1106, 2009.

[21] 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.

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

[23] X. Mao, *Stochastic Differential Equation and Applications*, Horwood Publishing Limited, England, UK, 1997.

[24] R. J. Hodrick and E. C. Prescott, "Postwar U.S business cycles: an empirical investigation," *Journal of Money, Credit and Banking*, vol. 29, no. 1, pp. 1–16, 1997.