Article

Two Stochastic Differential Equations for Modeling Oscillabolastic-Type Behavior

Antonio Barrera [1,†] **, Patricia Román-Román** [2,3,†] **and Francisco Torres-Ruiz** [2,3,*,†]

1   Departamento de Análisis Matemático, Estadística e Investigación Operativa y Matemática Aplicada,
     Facultad de Ciencias, Universidad de Málaga, Bulevar Louis Pasteur, 31, 29010 Málaga, Spain;
     antonio.barrera@uma.es
2   Departamento de Estadística e Investigación Operativa, Facultad de Ciencias, Universidad de Granada,
     Avenida Fuente Nueva s/n, 18071 Granada, Spain; proman@ugr.es
3   Instituto de Matemáticas de la Universidad de Granada (IEMath-GR), Calle Ventanilla 11,
     18001 Granada, Spain
*   Correspondence: fdeasis@ugr.es; Tel.: +34-958241000 (ext. 20056)
†   These authors contributed equally to this work.

check for updates

Abstract: Stochastic models based on deterministic ones play an important role in the description of growth phenomena. In particular, models showing oscillatory behavior are suitable for modeling phenomena in several application areas, among which the field of biomedicine stands out. The oscillabolastic growth curve is an example of such oscillatory models. In this work, two stochastic models based on diffusion processes related to the oscillabolastic curve are proposed. Each of them is the solution of a stochastic differential equation obtained by modifying, in a different way, the original ordinary differential equation giving rise to the curve. After obtaining the distributions of the processes, the problem of estimating the parameters is analyzed by means of the maximum likelihood method. Due to the parametric structure of the processes, the resulting systems of equations are quite complex and require numerical methods for their resolution. The problem of obtaining initial solutions is addressed and a strategy is established for this purpose. Finally, a simulation study is carried out.

Keywords: diffusion processes; growth model; oscillabolastic curve; stochastic differential equations

1. Introduction

Ordinary differential equations are one of the most frequently used mathematical tools for modeling and describing dynamical systems in general and growth phenomena in particular. In such models a quantitative variable evolves with time according to certain parameters and deterministic functions, and for their properties they are commonly applied to several fields of research.

Studying such models requires, on the one hand, understanding the evolving mechanisms of dynamical systems to properly explain their behavior. On the other, we must be able to include in them external information different from that supplied by the variables under study. In doing so we modify their behavior and exert external control on the evolution of the phenomenon under consideration.

However, in many situations, the resulting models do not adequately reproduce the observed phenomenon. The main reason is that deterministic models do not take into account certain disturbances that can greatly influence the final behavior. These disturbances may have multiple and various origins, which are not always quantifiable or may even be unknown. As stated by Li et al. [1], "*noises are ubiquitous in both nature and human society, such as fire, earthquake, climate warming, financial crisis, etc.*". Many applications require the modeling of external elements whose nature can be considered

stochastic, either because of the impossibility of making measurements or because said measurements present a high degree of uncertainty. It is, therefore, necessary to modify the deterministic model to introduce a random element.

Faced with that need, the design of models is obtained from the ordinary differential equation for the deterministic model, in which a random element is introduced, usually a function of a white noise. The result are stochastic models. Such modification requires considering stochastic differential equations (SDE) whose solutions are, under some conditions, diffusion processes. Not only does the use of these stochastic models provide a more realistic explanation of the variables under study, but it also allows us to study other important aspects (like inference or certain time variables that represent some features of the phenomenon).

As previously mentioned, ordinary differential equations are modified to give rise to a certain curve by introducing in them a white noise function. The choice of such function (usually motivated by specific characteristics of the studied phenomenon) can lead to a variety of stochastic processes, although they are related to the same deterministic curve. In addition, that choice can determine our specific approach to the stochastic solution process. In fact, even if the SDE has a solution (and sometimes it is not an explicit one), the resulting diffusion process is difficult to handle because the associated Kolmogorov equations have not yet been solved. For example, Feller pioneered this methodology by applying it to the case of the logistic equation and introducing an additional error term that was directly proportional to the square root of the variable. However, the backward and forward equations associated with the diffusion process solution of the SDE have never been solved (see Schurz [2] for other possible choices of the error term and associated diffusion processes). Furthermore, and regarding the logistic case, in the few cases where the transition density has been calculated (see Tuckwell [3]), it was impossible to find an explicit expression for certain important characteristics such as the mean of the process. This has led to theoretical works that analyze the existence of a solution for the SDE, study its stability properties (see Sun and Wang [4]), or propose simulation techniques (see Hu and Wang [5]). These considerations can be extended to other types of situations beyond the logistic case.

In light of the above, efforts have been made to address SDEs whose solutions are more manageable stochastic processes and also reproduce a certain pattern of observed behavior, i.e., their mean follows a given growth curve. This allows certain issues to be directly addressed, including inference or the study of temporal variables (such as first-passage-times), without having to resort to approximations such as those derived from the discretization of the SDE. The modification of certain ordinary differential equations by adding a multiplicative noise has given rise to processes whose mean is a growth curve. In this sense it is worth highlighting the role of the lognormal process with exogenous factors. As a matter of fact, concrete choices of such exogenous factors lead to processes by which we may study patterns of behavior modeled by a wide range of growth curves (see Román-Román et al. [6,7] and references therein). These references are related to the one-dimensional case, although there are extensions to multidimensionality such as the one proposed by Rupšys in [8], where a 4-variate Bertalanffy-type SDE is considered.

Other lines of action within the scope of SDEs should be mentioned. For example, the inclusion of delays in the modeling of certain phenomena (see Rupšys [9], Longtin [10], Sakthivel et al. [11]), the introduction of fractional Brownian motion and fractional calculus (Dung [12]), or the consideration of both methodologies at once (Moghaddam et al. [13]).

We must keep in mind that the ultimate objective of studying models such as these is modeling behavior patterns in real situations. Beyond the widely discussed classic growth curves (logistic, Gompertz, Bertalanffy, Richards, etc.), other interesting families of curves have been considered in recent times. Among them, the hyperbolastic functions of type I, II and III, introduced by Tabatabai et al. [14], stand out. The main feature of this type of curves is the introduction of hyperbolic functions into known models, thus revealing mobile inflection points and increasing the capability of the models to fit real data. Some works have proved the usefulness of such curves in the description of

dynamical phenomena. For example, Eby et al. [15] used them to study the growth of the solid Ehrlich carcinoma under particular treatments, obtaining a more accurate representation than those yielded by other classic curves such as Gompertz or Weibull. Tabatabai et al. [16] used the hyperbolastic type III model to describe the behavior of embryonic stem cells. Later, other models derived from hyperbolastic curves have been introduced, specifically the oscillabolastic and the T-model (Tabatabai et al. [17,18]). Focusing on the first one, the oscillabolastic curve was introduced as an appropriate model to describe dynamic phenomena that present oscillatory behaviors. The introduction of parameter-dependent trigonometric functions allows for oscillatory dynamics to be incorporated into the model and regulates both the amplitude and frequency of such oscillations. A progressive reduction in amplitude guarantees the asymptotic stability of the oscillatory behavior. In that paper the authors applied the oscillating model to two data sets: the first dealing with the oscillabolastic modeling of Ehrlich ascites tumor cells, and the second with the mean signal intensity of Hes1 gene expression in response to serum stimulation.

The main objective of this article is to present two diffusion processes associated with the oscillabolastic curve. Each will be obtained from a differential equation verified by the curve, and include an additive and a multiplicative error term, respectively. The rest of the paper is organized as follows: in Section 2 the oscillabolastic deterministic model is introduced and studied from a classical as well as an alternative differential formulation. The adaptation of these two formulations to stochastic models takes place in Section 3. To provide practical instances of application to the models presented, Section 4 addresses the maximum likelihood estimation of the parameters of each model. The complexity of the models and their high number of parameters make the resolution of the likelihood equations somewhat difficult. The use of numerical methods will thus require the prior estimation of initial solutions for the system, which is why a methodology for its calculation is suggested at the end of that section. Section 5 is a simulation study for both models that includes the generation of data for each process and the application of our proposed estimation methodology.

2. The Oscillabolastic Curve

The original formulation of the oscillabolastic curve is given by the function $u : \mathbb{R} \to \mathbb{R}$,

$$u(t) = \kappa + \alpha \sin(\beta t)/t + \gamma \operatorname{arcsinh}(\theta t)/t, \tag{1}$$

where $\kappa, \alpha, \beta, \gamma$ and θ are real parameters. Clearly, the function $u(t)$ is defined for all t and $\lim_{t \to \infty} u(t) = \kappa$. As easily seen, for $t \to 0^+$ one has $u(t) = \kappa + \alpha\beta + \gamma\theta$, so one can define $u(t)$ also at $t = 0$. However, in the context of growth phenomena, t is a time index and therefore we consider the curve defined in $[t_0, +\infty)$, $t_0 \geq 0$.

To study the behavior of (1), we start by observing that its expression is determined by two terms, say $u(t) = u_1(t) + u_2(t)$, where $u_1(t) = \kappa + \alpha \sin(\beta t)/t$ and $u_2(t) = \gamma \operatorname{arcsinh}(\theta t)/t$. It can therefore be seen that the behavior of the curve is governed mainly by $u_1(t)$. This function summarizes the oscillatory trend of the model, as well as its evolution towards κ, which we consider to be the population limit in a growth model (and not the maximum sustainable capacity that is typical of monotonous and positive growth models). Parameter α is responsible for the amplitude of the oscillations, which become larger as $|\alpha|$ grows. Furthermore, their frequency is directly proportional to $|\beta|$.

The flexibility of the oscillabolastic curve is clearly increased thanks to the influence of $u_2(t)$. In this case, the hyperbolic arcsine allows $u_1(t)$ to be deformed, thus making the oscillatory pattern evolve along a curved path. This in turn has an impact on the sequence of local extremes of $u(t)$ (local minima and maxima at each oscillation), since one of the functions ($u(t)$ or $u_1(t)$) will present a wider range at initial values than the other, which will remain closer to constant value κ. In any case, both will eventually converge to κ. Figure 1 shows the two parts of an oscillabolastic curve for $\kappa = 1, \alpha = -1, \beta = 1, \gamma = -1$ and $\theta = 1$, as well as the resulting curve $u(t)$. It can be seen how $u_2(t)$ influences the final curve by modifying the behavior of $u_1(t)$. This interaction is indicated by a

dashed line representing translated function $u_2(t) + v$, where v is the average difference between $u(t)$ and $u_2(t)$.
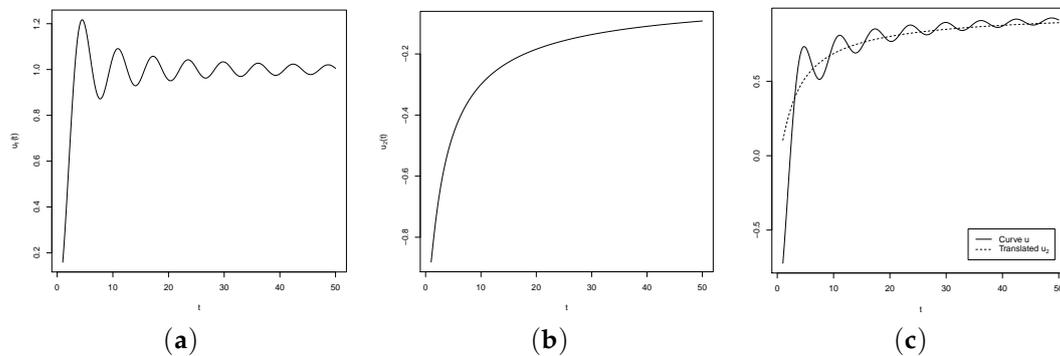
Figure 1. $u_1(t)$ curve (**a**), $u_2(t)$ curve (**b**), and oscillabolastic $u(t)$ curve together with $u_2(t)$ (dashed line) moved to indicate the trend (**c**). The translation of $u_2(t)$ is the average between the difference $u(t) - u_2(t)$ for every t. Parameters: $\kappa = 1, \alpha = -1, \beta = 1, \gamma = -1, \theta = 1, t_0 = 1$.

Another fundamental aspect of $u_2(t)$ is its ability to make the final curve smoother, and to partially compensate the amplitude of oscillations from $u_1(t)$. In this sense, high values of γ/α cause a deformation of the oscillations, increasing the difference between the increasing and decreasing phases.

In any case, the behavior of the oscillabolastic curve $u(t)$ is characterized by four parameters. It is not easy to establish a clear pattern of influence for each without taking into account the others, as well as the time interval in which the curve is evaluated.

An important characteristic of the oscillabolastic curve is the presence of multiple inflections, a logical consequence of the oscillations of the model. Equation $d^2u(t)/dt^2 = 0$ leads to the following equation, verified for any inflection time t_*,

$$2\gamma \operatorname{arcsinh}(\theta t_*) - \frac{\gamma \theta t_* (2 + 3\theta^2 t_*^2)}{(1 + \theta^2 t_*^2)^{3/2}} = 2\alpha\beta t_* \cos(\beta t_*) + \alpha \sin(\beta t_*)(\beta^2 t_*^2 - 2). \tag{2}$$

From an analytical approach, the function $u(t)$ satisfies, under the initial condition $u(t_0) = u_0 \in \mathbb{R}$, the linear ordinary differential equation

$$\frac{d}{dt}u(t) = (g_\delta(t) - u(t)) t^{-1}, \tag{3}$$

for $t \geq t_0 > 0$, being

$$g_\delta(t) = \kappa + \alpha\beta \cos(\beta t) + \gamma \theta \left(1 + \theta^2 t^2\right)^{-1/2} \tag{4}$$

a real function defined in the same domain as $u(t)$ and where δ denotes the parametric vector $\delta = (\kappa, \alpha, \beta, \gamma, \theta)^T$.

On the other hand, it is possible to reformulate the oscillabolastic curve as

$$u(t) = \kappa \left(1 + \alpha' \sin(\beta t)/t + \gamma' \operatorname{arcsinh}(\theta t)/t\right),$$

where the relationship between the parameters of both formulations would be given by $\alpha' = \alpha/\kappa$ and $\gamma' = \gamma/\kappa$.

In this case, it can be verified that the reformulated curve is the solution to the linear ordinary differential equation

$$\frac{d}{dt}u(t) = h_{\delta'}(t)\, u(t), \qquad t \geq t_0 > 0, \tag{5}$$

with initial condition $u(t_0) = u_0$, and where function $h_{\delta'}(t)$ is expressed as

$$h_{\delta'}(t) = \frac{\alpha' \left(\beta t \cos(\beta t) - \sin(\beta t)\right) + \gamma' \left(\frac{\theta t}{\sqrt{1+\theta^2 t^2}} - \text{arcsinh}(\theta t)\right)}{t^2 + \alpha' t \sin(\beta t) + \gamma' t \, \text{arcsinh}(\theta t)}, \qquad t \geq t_0 > 0, \qquad (6)$$

being $\delta' = (\alpha', \beta, \gamma', \theta)^T$.

Both differential equations, (3) and (5), define deterministic models of evolution for the oscillabolastic curve. The most notable difference between them is their structure: in the first case we have a non-homogeneous linear differential equation with non-constant coefficients, while in the second it is reduced to a multiplicative model.

Given the potential applications of both models, their stochastic extensions will be addressed in the following sections. The main idea is to make the mean of the resulting stochastic process able to describe oscillabolastic-type behaviors, which opens the door for observed patterns of this type to be reproduced.

3. Oscillabolastic-Type Diffusion Processes

This section introduces stochastic extensions for the deterministic oscillabolastic model in light of the two differential equations described in the previous section. The starting points will be ordinary differential Equations (3) and (5). Although both are linear differential equations, the diffusion processes to which they give rise have different characteristics. The former is a Gaussian process (of the non-homogeneous Ornstein–Uhlenbeck type), while the second is related to the non-homogeneous geometric Brownian motion, or lognormal process with exogenous factors.

3.1. Oscillabolastic Gaussian Diffusion Process

This process is obtained by adding a random component to (3). Specifically, a white noise with variance σ^2 is included in the model, where $\sigma > 0$ is a parameter that marks the random influence on the model. Writing the resulting equation in the traditional form of stochastic differential equations, we have

$$dX^G(t) = \left(g_\delta(t) - X^G(t)\right) t^{-1} dt + \sigma \, dW(t), \qquad (7)$$

where $\{W(t); t \geq t_0\}$ is the standard Wiener process in $\mathbb{R}$, which we consider independent from the initial variable $X^G(t_0)$, whereas g_δ is given by (4).

Equation (7) is a linear SDE verifying the conditions of existence and uniqueness of solution. In fact, functions $t \mapsto -t^{-1}$ and $t \mapsto g_\delta(t)t^{-1}$ are measurable and bounded in every subinterval $[t_0, T]$, with $T > t_0$, of $[t_0, +\infty)$, which guarantees the existence and uniqueness of a global solution (i.e., defined in $[t_0, +\infty)$) for each initial condition $X_0^G = X^G(t_0)$.

This solution is a stochastic diffusion process $\{X^G(t); t \geq t_0\}$, adapted to the natural filtration generated by the Borel sets in one dimension and characterized by drift $(g_\delta(t) - x)/t$ and constant infinitesimal variance σ^2. The explicit formulation of the process is given in terms of the solution of ordinary differential equation $dx(t)/dt = -x(t)/t$, leading to

$$X^G(t) = \left(t_0 X_0^G + \int_{t_0}^{t} g_\delta(s) ds + \sigma \int_{t_0}^{t} s \, dW(s)\right) t^{-1}.$$

Initial condition X_0^G is fundamental to establishing the nature of the process. Indeed, if X_0^G is a constant (degenerate) or a Gaussian random variable, then the solution of (7) is a Gaussian diffusion process (from now on we will consider both scenarios). In addition, in the case where X_0^G is constant, the independence of the increments of the process is also guaranteed.

On the other hand, condition $E\left[|X_0^G|^p\right] < \infty$ guarantees the existence of the p-th order moment for $X^G(t)$. Specifically, the expressions of the mean and covariance functions are respectively

$$m_\delta(t) = E[X^G(t)] = \left(t_0 E[X_0^G] + \kappa\Delta_0 t + \alpha\Delta_0 \sin(\beta t) + \gamma\Delta_0 \operatorname{arcsinh}(\theta t)\right) t^{-1},$$

$$C(t,s) = \operatorname{Cov}(X^G(t), X^G(s)) = (\sigma_0 t_0)^2 / (st) + \sigma^2 \left(t^3 \wedge s^3 - t_0^3\right) / (3st),$$

for $s, t \in [t_0, +\infty)$, being $\sigma_0^2 = \operatorname{Var}[X_0^G]$. Here Δ_0 denotes increasing operator $\Delta_0 f(t) = f(t) - f(t_0)$ for a function f, whereas $t \wedge s = Min(s, t)$.

Regarding the mean function $m_\delta(t)$, note that it can be expressed as

$$m_\delta(t) = u(t) + \frac{t_0}{t}\left(E[X_0^G] - u_0\right),$$

where $u_0 = u(t_0)$ is the initial value of the oscillablastic curve. Also note that when constructing the stochastic model, the main goal is for the mean function of the process to coincide with that of a prescribed oscillablastic model, i.e., $m_\delta(t) = u(t)$ for all t. Regarding the last expression, this condition holds when $E[X_0^G] = u_0$. However, the asymptotic tendency of the process (the behavior of its mean at the limit) is similar to that of the oscillablastic deterministic model. Indeed, the difference between the mean and the curve, for each time instant, decreases as t tends to ∞, i.e.,

$$\lim_{t\to\infty} |m_\delta(t) - u(t)| = 0.$$

With regard to the transition density function, which is required to implement the inference procedures that will be described in the next section, for any time instants $t > s \geq t_0$ the conditioned variable $X^G(t)|X^G(s) = y$ is normal, with mean and variance given respectively by expressions

$$\mu_\delta(s,t) = m_\delta(t) + \sigma(t)\rho(s,t)(y - m_\delta(s))/\sigma(s),$$

$$\nu(s,t)^2 = \left(1 - \rho(s,t)^2\right)\sigma(t)^2,$$

where $\sigma(t) = C(t,t)^{1/2}$ is the standard deviation of $X^G(t)$ (resp. for s) and $\rho(s,t) = C(s,t)/(\sigma(s)\sigma(t))$ is the correlation function. Both expressions can also be written in the form

$$\mu_\delta(s,t) = \frac{1}{t}\left(sy + \int_s^t g_\delta(r)dr\right)$$

and

$$\nu(s,t)^2 = \frac{\sigma^2}{3}\left(\frac{t^3 - s^3}{t^2}\right),$$

from which the probability transition density function is given by

$$f(x,t|y,s) = \left(2\pi\frac{\sigma^2}{3}\left(\frac{t^3 - s^3}{t^2}\right)\right)^{-\frac{1}{2}} \exp\left(-\frac{\left(x - \frac{s}{t}y - \frac{1}{t}\int_s^t g_\delta(r)dr\right)^2}{2\frac{\sigma^2}{3}\left(\frac{t^3-s^3}{t^2}\right)}\right). \tag{8}$$

3.2. Oscillablastic Diffusion Process Derived from the Non-Homogeneous Lognormal Diffusion Process

From Equation (5), it is possible to obtain a stochastic model whose mean function coincides with the original deterministic model, i.e., the oscillablastic curve. As a matter of fact, this differential equation represents a Malthusian model with a time-dependent fertility $h_{\delta'}(t)$ given by (6).

The introduction of a white noise $\varrho(t)$ with variance σ^2 in this fertility component, accomplished by replacing $h_{\delta'}(t)$ with $h_{\delta'}(t) + \varrho(t)$, leads to the SDE

$$dX^L(t) = h_{\delta'}(t)X^L(t)\,dt + \sigma X^L(t)\,dW(t), \qquad X^L(t_0) = X_0^L, \tag{9}$$

where $W(t)$ is the standard Wiener process, independent from X_0^L, $t \geq t_0$. Please note that the reformulation of the original curve leads to a model with one parameter less than the Gaussian model.

In this case, the existence and uniqueness of the solution is ensured by the continuity of the $h_{\delta'}(t)$ function. Thus, the solution of (9) is a stochastic diffusion process $\{X^L(t); t \geq t_0\}$ characterized by drift $h_{\delta'}(t)x$ and infinitesimal variance $\sigma^2\,x^2$. Furthermore, a closed-form expression for the solution is given, for $t \geq t_0$, by

$$X^L(t) = X_0^L \exp\left(H_{\xi}(t_0, t) + \sigma(W(t) - W(t_0))\right),$$

where

$$H_{\xi}(s, t) = \int_s^t h_{\delta'}(u)du - \frac{\sigma^2}{2}(t - s) = \log\frac{Q_{\delta'}(t)}{Q_{\delta'}(s)} - \frac{\sigma^2}{2}(t - s), \qquad s < t,\ \boldsymbol{\xi} = \left(\delta'^T, \sigma^2\right)^T,$$

being

$$Q_{\delta'}(t) = 1 + \alpha'\sin(\beta t)/t + \gamma'\operatorname{arcsinh}(\theta t)/t.$$

This process is a particular case of the non-homogeneous lognormal diffusion, or lognormal diffusion process with exogenous factors. The inclusion of the $h_{\delta'}(t)$ function in the drift justifies the terminology: the infinitesimal mean of the process can be affected by the inclusion of time-dependent external factors.

As regards the distribution of the process, if X_0^L is distributed according to a lognormal distribution $\Lambda_1\left[\mu_0; \sigma_0^2\right]$ or is a degenerate random variable ($\mathrm{P}[X_0^L = x_0] = 1$), all the finite-dimensional distributions of the process are lognormal. Concretely, $\forall n \in \mathbb{N}$ and $t_1 < \cdots < t_n$, vector $(X^L(t_1), \ldots, X^L(t_n))^T$ is distributed according to a n-dimensional lognormal distribution $\Lambda_n[\boldsymbol{\varepsilon}, \boldsymbol{\Sigma}]$, where the components of the vector $\boldsymbol{\varepsilon}$ are $\varepsilon_i = \mu_0 + H_{\xi}(t_0, t_i)$, $i = 1, \ldots, n$, being $\sigma_{ij} = \sigma_0^2 + \sigma^2(\min(t_i, t_j) - t_0)$, $i, j = 1, \ldots, n$, those of the matrix $\boldsymbol{\Sigma}$.

As far as the transition probability is concerned, it is obtained from that of $(X^L(s), X^L(t))^T$, $s < t$. Concretely,

$$X^L(t) \mid X^L(s) = y \rightsquigarrow \Lambda_1\left(\log y + H_{\xi}(s, t), \sigma^2(t - s)\right), \qquad s < t,$$

whose density function is

$$f(x, t|y, s) = \frac{1}{x\sqrt{2\pi\sigma^2(t - s)}}\exp\left(-\frac{\log\frac{x}{y} - H_{\xi}(s, t)}{2\sigma^2(t - s)}\right). \tag{10}$$

Once the distribution of the process has been established, we may calculate a variety of its associated characteristics including the mean and conditioned mean functions, whose expressions are

$$m_{\delta'}(t) = E[X_0^L]\frac{Q_{\delta'}(t)}{Q_{\delta'}(t_0)}$$

and

$$m_{\delta'}(t|t_0) = E[X^L(t)|X^L(t_0) = x_0] = x_0\exp\left(H_{\xi}(t_0, t) + \frac{\sigma^2}{2}\right) = x_0\frac{Q_{\delta'}(t)}{Q_{\delta'}(t_0)},$$

respectively. Both functions are oscillabolastic curves of the type introduced in the previous section. Other characteristics of the process can be obtained from the results included in Román et al. [6].

4. Estimation of the Parameters

The main useful application of the diffusion processes described in the previous section is the description of dynamic phenomena subject to random influences, in particular growth phenomena. For that reason, the estimation of the parameters included in the models is especially relevant. Among the procedures available we will employ the maximum likelihood estimation method. The estimation of the parameters involves solving a system of equations, which will require using adequate numerical techniques.

Let us consider d sample paths observed at time instants t_{ij}, $(i = 1, \ldots, d, j = 1, \ldots, n_i)$. Please note that neither the sample sizes nor the times of observation have to be the same, although we will suppose that the first time of observation is common for all sample paths, i.e., $t_{i1} = t_0$, $i = 1, \ldots, d$. Let $\mathbf{X}_i^T$ be the vector containing the random variables of the i-th sample-path, i.e., $\mathbf{X}_i = (X(t_{i1}), \ldots, X(t_{i,n_i}))^T$, $i = 1, \ldots, d$, and denote $\mathbf{X} = (\mathbf{X}_1^T | \cdots | \mathbf{X}_d^T)^T$.

For simplicity, we consider a degenerate initial distribution, i.e., $P[X_0 = x_0] = 1$. The reason is that the parameters of the distribution in the non-degenerate case (normal or lognormal depending on the process) are usually considered functionally independent from the parameters involved in the transition distribution, so the estimation of these last parameters is not influenced by the estimation of the others. In any case, if X_0 were not degenerate its parameters could not be estimated from the values of the sample paths at the first time instant of observation.

In what follows, and for the sake of simplicity in the notation, we will not differentiate between the processes when noting the observed values. Obviously, the context determines the meaning in each case.

4.1. The Gaussian Case

Let us consider a fixed value $\mathbf{x}$ of $\mathbf{X}$. From (8) the likelihood function can be obtained, whose logarithm is

$$\log L_{\mathbf{x}}(\boldsymbol{\xi}) = -\frac{n}{2} \log \frac{2\pi}{3} - \frac{n}{2} \log \sigma^2 - \frac{1}{2} \sum_{i=1}^{d} \sum_{j=2}^{n_i} \log \frac{\Delta t_{ij}^3}{t_{ij}^2} - \frac{3}{2\sigma^2} \sum_{i=1}^{d} \sum_{j=2}^{n_i} \frac{1}{\Delta t_{ij}^3} \left(\Delta \left(t_{ij} x_{ij} \right) - G_{ij}^{\delta} \right)^2,$$

where $n = \sum_{i=1}^{d}(n_i - 1)$ and $G_{ij}^{\delta} := \int_{t_{ij-1}}^{t_{ij}} g_{\delta}(s)ds$. Now operator Δ (applied to a generic function f) denotes the increments $\Delta f(t_{ij}) = f(t_{ij}) - f(t_{ij-1})$.

To obtain the maximum likelihood estimates, we must calculate the derivatives of $\log L_{\mathbf{x}}(\boldsymbol{\xi})$ with respect to the parameters. In this sense, the derivatives with respect to the components of δ are

$$\partial_\tau \log L_{\mathbf{x}}(\boldsymbol{\xi}) = \frac{3}{\sigma^2} \sum_{i=1}^{d} \sum_{j=2}^{n_i} \frac{1}{\Delta t_{ij}^3} \left(\Delta \left(t_{ij} x_{ij} \right) - G_{ij}^{\delta} \right) \partial_\tau G_{ij}^{\delta}, \qquad \tau \in \delta \tag{11}$$

where

$$\partial_\tau G_{ij}^{\delta} = \begin{cases} \Delta t_{ij} & \text{if } \tau = \kappa, \\[1.5em] \Delta \sin(\beta t_{ij}) & \text{if } \tau = \alpha, \\[1.5em] \alpha \, \Delta \left(t_{ij} \cos(\beta t_{ij}) \right) & \text{if } \tau = \beta, \\[1.5em] \Delta \operatorname{arcsinh}(\theta t_{ij}) & \text{if } \tau = \gamma, \\[1.5em] \gamma \, \Delta \left(t_{ij}^2 \Big/ \sqrt{1 + \theta^2 t_{ij}^2} \right) & \text{if } \tau = \theta. \end{cases}$$

On the other hand, the partial derivative with respect to σ^2 is

$$\partial_{\sigma^2} \log L_{\mathbf{x}}(\boldsymbol{\xi}) = -\frac{n}{2\sigma^2} + \frac{3}{2\left(\sigma^2\right)^2} \sum_{i=1}^{d} \sum_{j=2}^{n_i} \frac{1}{\Delta t_{ij}^3} \left(\Delta\left(t_{ij}x_{ij}\right) - G_{ij}^{\delta}\right)^2. \tag{12}$$

Finally, by using (11) and (12), and equating to zero the partial derivative of the loglikelihood, one gets:

$$\sum_{i=1}^{d} \sum_{j=2}^{n_i} \frac{1}{\Delta t_{ij}^3} \left(\Delta\left(t_{ij}x_{ij}\right) - k\Delta t_{ij} - \alpha\Delta\sin(\beta t_{ij}) - \gamma\Delta\operatorname{arcsinh}(\theta t_{ij})\right)\Delta t_{ij} = 0,$$

$$\sum_{i=1}^{d} \sum_{j=2}^{n_i} \frac{1}{\Delta t_{ij}^3} \left(\Delta\left(t_{ij}x_{ij}\right) - k\Delta t_{ij} - \alpha\Delta\sin(\beta t_{ij}) - \gamma\Delta\operatorname{arcsinh}(\theta t_{ij})\right)\Delta\sin(\beta t_{ij}) = 0,$$

$$\sum_{i=1}^{d} \sum_{j=2}^{n_i} \frac{1}{\Delta t_{ij}^3} \left(\Delta\left(t_{ij}x_{ij}\right) - k\Delta t_{ij} - \alpha\Delta\sin(\beta t_{ij}) - \gamma\Delta\operatorname{arcsinh}(\theta t_{ij})\right)\alpha\Delta\left(t_{ij}\cos(\beta t_{ij})\right) = 0,$$

$$\sum_{i=1}^{d} \sum_{j=2}^{n_i} \frac{1}{\Delta t_{ij}^3} \left(\Delta\left(t_{ij}x_{ij}\right) - k\Delta t_{ij} - \alpha\Delta\sin(\beta t_{ij}) - \gamma\Delta\operatorname{arcsinh}(\theta t_{ij})\right)\Delta\operatorname{arcsinh}(\theta t_{ij}) = 0, \tag{13}$$

$$\sum_{i=1}^{d} \sum_{j=2}^{n_i} \frac{1}{\Delta t_{ij}^3} \left(\Delta\left(t_{ij}x_{ij}\right) - k\Delta t_{ij} - \alpha\Delta\sin(\beta t_{ij}) - \gamma\Delta\operatorname{arcsinh}(\theta t_{ij})\right)\gamma\Delta\left(t_{ij}^2 \Big/ \sqrt{1 + \theta^2 t_{ij}^2}\right) = 0,$$

$$\frac{3}{n\sigma^2} \sum_{i=1}^{d} \sum_{j=2}^{n_i} \frac{1}{\Delta t_{ij}^3} \left(\Delta\left(t_{ij}x_{ij}\right) - k\Delta t_{ij} - \alpha\Delta\sin(\beta t_{ij}) - \gamma\Delta\operatorname{arcsinh}(\theta t_{ij})\right)^2 - 1 = 0.$$

The resolution of this system will be determined by the application of numerical techniques. Such procedures will require a set of initial solutions. This problem, common to the two models presented, will be addressed in Section 4.3.

4.2. Lognormal Case

Regarding oscillabolastic process $X^L(t)$, given by SDE (9), the maximum likelihood estimation can be addressed by means of the procedure described in [6]. This approach has already been used succesfully in the inferential treatment of the Gompertz multisigmoidal diffusion process [7], as well as in the generalization of the classic Weibull model, in particular for the hyperbolastic type III diffusion process [19].

In this case, and considering probability transition density function (10), the logarithm of the likelihood function is given by

$$\log L_{\mathbf{x}}(\boldsymbol{\xi}) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{Z_1 + \Phi_{\boldsymbol{\xi}} - 2\Gamma_{\boldsymbol{\xi}}}{2\sigma^2},$$

where

$$Z_1 = \sum_{i=1}^{d} \sum_{j=2}^{n_i} \frac{1}{\Delta t_{ij}} \log\frac{x_{ij}}{x_{ij-1}}, \quad \Phi_{\boldsymbol{\xi}} = \sum_{i=1}^{d} \sum_{j=2}^{n_i} \frac{1}{\Delta t_{ij}} \left(m_{ij}^{\boldsymbol{\xi}}\right)^2, \quad \Gamma_{\boldsymbol{\xi}} = \sum_{i=1}^{d} \sum_{j=2}^{n_i} \frac{m_{ij}^{\boldsymbol{\xi}}}{\Delta t_{ij}} \log\frac{x_{ij}}{x_{ij-1}}$$

being $m_{ij}^{\boldsymbol{\xi}} = H_{\boldsymbol{\xi}}(t_{ij-1}, t_{ij})$.

The estimate of vector $\boldsymbol{\xi}$ (see [6] for details) follows by solving the system of equations

$$\Psi_{\delta'} - \Omega_{\boldsymbol{\xi}} = 0, \tag{14}$$

$$Z_1 + \Phi_{\boldsymbol{\xi}} - 2\Gamma_{\boldsymbol{\xi}} - \sigma^2 Z_2 + \sigma^2 Y_{\boldsymbol{\xi}} = n\sigma^2, \tag{15}$$

where

$$\Omega_\xi = \frac{1}{2}\frac{\partial \Phi_\xi}{\partial \delta'^T}, \quad \Psi_{\delta'} = \frac{\partial \Gamma_\xi}{\partial \delta'^T}, \quad Y_\xi = -\frac{\partial \Phi_\xi}{\partial \sigma^2}, \quad Z_2 = -2\frac{\partial \Gamma_\xi}{\partial \sigma^2}.$$

These functions are obtained from the partial derivatives of the m_{ij}^ξ functions:

$$\tau\omega_{\delta'} = \frac{\partial m_{ij}^\xi}{\partial \tau} = \frac{1}{Q_{\delta'}(t_{ij})}\left.\frac{\partial Q_{\delta'}(t)}{\partial \tau}\right|_{t=t_{ij}} - \frac{1}{Q_{\delta'}(t_{ij-1})}\left.\frac{\partial Q_{\delta'}(t)}{\partial \tau}\right|_{t=t_{ij-1}}, \quad \tau \in \delta',$$

$$\frac{\partial m_{ij}^\xi}{\partial \sigma^2} = -\frac{1}{2}\Delta t_{ij},$$

where

$$\frac{\partial Q_{\delta'}}{\partial \tau} = \begin{cases} \sin(\beta t)/t & \text{if } \tau = \alpha', \\[2ex] \alpha'\cos(\beta t) & \text{if } \tau = \beta, \\[2ex] \operatorname{arcsinh}(\theta t)/t & \text{if } \tau = \gamma', \\[2ex] \gamma'/\sqrt{1+\theta^2 t^2} & \text{if } \tau = \theta. \end{cases}$$

From these last expressions, and from those of Ω_ξ and $\Psi_{\delta'}$, the subsystem of Equation (14) can be written in the form

$$\tau\Xi_{\delta'} + \frac{\sigma^2}{2}\,\tau\overline{\Omega}_{\delta'} = 0, \; \tau \in \delta', \tag{16}$$

with

$$\tau\Xi_{\delta'} = \sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{\log\left(x_{i,j}/x_{i,j-1}\right) - \log\left(Q_{\delta'}(t_{ij})/Q_{\delta'}(t_{i,j-1})\right)}{\Delta t_{ij}}\,\tau\omega_{\delta'}^{i,j,j-1}, \quad \tau\overline{\Omega}_{\delta'} = \sum_{i=1}^{d}\sum_{j=2}^{n_i}\tau\omega_{\delta'}^{i,j,j-1}, \; \tau \in \delta'.$$

On the other hand, and after some calculus, Equation (15) transforms into

$$\sigma^2\left[n + \sigma^2 Z_3/4\right] + 2X_1^{\delta'} - X_2^{\delta'} - Z_1 = 0, \tag{17}$$

being

$$X_1^{\delta'} = \sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{\left(\log\left(Q_{\delta'}(t_{ij})/Q_{\delta'}(t_{i,j-1})\right)\right)^2}{\Delta t_{ij}}, \quad X_2^{\delta'} = \sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{\log\left(x_{i,j}/x_{i,j-1}\right)\log\left(Q_{\delta'}(t_{ij})/Q_{\delta'}(t_{i,j-1})\right)}{\Delta t_{ij}},$$

and $Z_3 = \sum_{i=1}^{d}\sum_{j=2}^{n_i}\Delta t_{ij}$.

As in the previous case, system of Equations (16) and (17) cannot be solved explicitly, and it is, therefore, necessary to use numerical methods such as Newton-Raphson's, for which an initial solution is required. In the next subsection we present a strategy to provide such a solution based on the information provided by the sample data, and applicable to the two resulting systems of equations.

4.3. Initial Solutions

As already mentioned, the systems of likelihood equations obtained in the preceding subsections have a high degree of complexity due to the high number of parameters involved (although in the lognormal case this number is reduced) and to the presence of random data. Their resolution, therefore, will require the use of numerical techniques. For such procedures to be applied a prior choice must be made regarding initial solutions.

　　This section proposes a strategy based on the information provided by the sample data. The key to the procedure focuses on the sample mean function, which is calculated from the values of the sample paths at each time instant of observation t_i. Since the procedure needs to evaluate this function over the entire observation interval, a spline function is later fitted to the mean values of the observed sample paths. Furthermore, the procedure is applicable to the two processes introduced earlier. The goal is to make the theoretical mean functions of both processes approximate the oscillabolastic curve (1), although in the Gaussian case this approach is asymptotic. We expose the strategy for the Gaussian case. For the lognormal-type process it is deduced by simply taking into account the relationship between the parameters which has been described in Section 2.

　　The following is the suggested procedure for each parameter:

- κ: This is the limit value of the model, and represents its asymptotic behavior when $t \to +\infty$. However, due to oscillations, this value is not the maximum (which usually occurs in monotonous positive growth models, giving rise to the *carrying capacity* of the system under consideration). Since the curve approaches κ, a possible option to obtain an initial value κ_0 is to take the last value of the sample mean function. However, this depends largely on the length of the time interval, as well as on the amplitude of the oscillations and/or of the curvature of the model. These effects can lead to estimates of κ far from its true value. For greater accuracy, an alternative approach would be to take the average value between the sequences of local extreme values (maxima and minima) of the sample mean function.

- $\alpha, \beta, \gamma, \theta$: The calculation of initial values for these parameters is quite complex. The starting point is to have an approximation for one of the instants of time (t_*) in which an inflection is observed in the mean of the process. To obtain it, a maximum of the derivative of the sample mean function is numerically approached.

　　On the other hand, it can be observed that (1) is a linear function in α and γ. Therefore, considering t_0 and t_*, and the value κ_0 previously calculated, it is possible to obtain an expression of both parameters as a function of β and θ from the system of equations

$$
\kappa_0 = u_0 - \alpha \sin(\beta t_0)/t_0 - \gamma \operatorname{arcsinh}(\theta t_0)/t_0,
$$
$$
\kappa_0 = u_* - \alpha \sin(\beta t_*)/t_* - \gamma \operatorname{arcsinh}(\theta t_*)/t_*. \tag{18}
$$

Please note that u_0 and u_* can be obtained from the spline function previously adjusted to the mean of the sample paths. We shall denote by $\alpha_{\beta,\theta}$ and $\gamma_{\beta,\theta}$ the resulting values. Now the problem has been reduced to finding initial values for β and θ. To this end we look for pairs of values (β, θ) within a two-dimensional bounded region satisfying the inflection condition (2) for a predefined error threshold, say ϵ, i.e., it must hold

$$
\left| \frac{\gamma_{\beta,\theta} \theta t_* (2 + 3\theta^2 t_*^2)}{(1 + \theta^2 t_*^2)^{3/2}} - 2\gamma_{\beta,\theta} \operatorname{arcsinh}(\theta t_*) + 2\alpha_{\beta,\theta} \beta t_* \cos(\beta t_*) + \alpha_{\beta,\theta} \sin(\beta t_*)(\beta^2 t_*^2 - 2) \right| < \epsilon,
$$

being t_* the inflection time instant previously obtained. We suggest using values of ϵ between 0.0001 and 0.1, depending on the order of magnitude of the sample data. Initial values β_0 and θ_0 will be the mean of the resulting values for each parameter, respectively. Finally, the initial values for α and γ are obtained from $\alpha_{\beta_0,\theta_0}$ and $\gamma_{\beta_0,\theta_0}$, respectively.

- σ^2: The initial value for σ^2 is obtained from the estimation of the variance of the one-dimensional distributions of the processes. For instance, in the case of the Gaussian oscillabolastic process, the variance is (please remember that we have considered a degenerate initial distribution)

$$
Var[X(t)] = \sigma^2 \frac{t^3 - t_0^3}{3t^2},
$$

so, if we denote by σ_i^2 the sample variance at t_i, a linear regression model of the σ_i^2 values versus t_i ($i = 1, \ldots, n$) is considered in order to obtain the estimated value for σ^2. It should be noted that the sample variance can sometimes present a rather volatile behavior. Therefore, in practice it is advisable to procure a smoothed version of said function before obtaining values for σ_i^2.

For the other oscillabolastic-type diffusion process, for which (see Section 3.2) $X(t) \rightsquigarrow \Lambda_1[\varepsilon, \varsigma]$ being $\varsigma = \sigma^2(t - t_0)$, we can avoid using the sample variance function. In point of fact, it is well-known that for a lognormal distribution $\Lambda_1[\varepsilon, \varsigma]$ the quotient between the arithmetic mean and the geometric one provides an estimation of ς. This leads, for each t_i, and when the initial distribution is degenerate, to an estimate of $\sigma^2(t_i - t_0)$ in the distribution of $X(t)$, say, $\sigma_i^2 = 2\log(m_i/m_i^g)$, $i = 1, \ldots$, where m_i and m_i^g are, respectively, the values of the mean and the geometric sample mean of the sample paths at t_i. Again, the initial value for σ^2 is obtained by performing a simple linear regression of σ_i^2 values against t_i.

Please note that if X_0 is not a degenerate random variable, σ_i^2 provides an estimate of $\sigma_0^2 + \sigma^2(t_i - t_0)$ for each t_i. In that case, σ_0^2 can be previously estimated from the values of the sample paths at t_0 and then its values introduced in the regression function.

5. Simulation Study

In this section, a simulation study illustrates the behavior of the oscillabolastic process, as well as the application of the proposed methodology for estimating the parameters of the processes under consideration. One of its most interesting aspects is how initial solutions are obtained for the subsequent estimation of the parameters of the process, to which end we have applied the methodology proposed in the previous subsection.

5.1. Simulation in the Gaussian Case

Data from the oscillabolastic Gaussian process are obtained by simulating 20 sample paths in the time interval $[1, 10]$, with degenerate initial distribution $x_0 = 1$. As for the values of the parameters, we have started by considering $\kappa = 2.5, \alpha = -2, \beta = 1.5, \gamma = 1$ and $\theta = 0.5$. The parameter corresponding to the infinitesimal variance is, in this case, $\sigma^2 = 0.0025$. Simulated paths are shown in Figure 2.
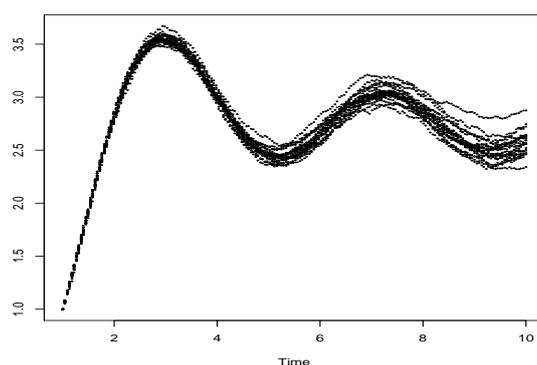
Figure 2. Simulated sample paths for the oscillabolastic Gaussian process. Parameters: $\kappa = 2.5$, $\alpha = -2, \beta = 1.5, \gamma = 1, \theta = 0.5$ and $\sigma^2 = 0.0025$.

Following the method described in Section 4.1, the system of maximum likelihood Equation (13) is solved by numerical procedures. To do this, initial solutions to the parameters are obtained by applying the methodology described in Section 4.3. The initial value for κ is selected by taking the last value of the observed mean, in this case $\kappa_0 = 2.58$. One inflection point is calculated from (2) after fitting a spline function to the observed mean values, resulting in $t_* = 1.38$, and being $u_* = 1.71$ the value of the curve at this point. With this information, initial values β_0 and θ_0 for β and θ, respectively, are found by

holding the inflection condition with an error threshold $\epsilon = 5 \times 10^{-4}$. The search region is determined by the square $[0.1, 3] \times [0.1, 1]$. In fact, the number of oscillations of the observed data suggests that β must lie between 0.1 and 3 (and might even be 2). On the other hand, the absence of a pronounced curvature implies low values of θ. Finally, the initial values are $\beta_0 = 1.5$ and $\theta_0 = 0.3$. From these values, and using the expressions $\alpha_{\beta_0,\theta_0}$ and $\gamma_{\beta_0,\theta_0}$ obtained from (18), we find the corresponding initial values for α and β, concretely $\alpha_0 = -1.97$ and $\gamma_0 = 1.1$. Finally, the initial value for σ^2 is obtained from the estimation of the slope of the regression line of observed variability against time, resulting in $\sigma_0^2 = 0.0029$.

Once the initial values of the parameters have been determined, they are used to solve the system of equations. Detailed results for estimated values of each parameter are shown in Table 1. Please note that in this case, initial values coincide with estimated ones, mainly due to low variability of the sample paths.

Table 1. Results for simulated Gaussian oscillabolastic diffusion.

	κ	α	β	γ	θ	σ^2
Estimated values	2.584	-1.973	1.500	1.1	0.300	0.0029

To evaluate the quality of the estimation of the process, we have considered the absolute relative errors between the sample mean of the simulated process and the estimated one, i.e.,

$$\text{RAE} = \frac{1}{n} \sum_{j=1}^{n} \frac{|m_j - \hat{m}_j|}{m_j},$$

where m_j and $\hat{m}_j$ are the values of the sample mean function and the ones estimated at t_j, $j = 1, \ldots, n$.

In this case, the value of this error is $\text{RAE} = 0.024$, which indicates a very good fit between the observed and estimated mean functions from the values found for the parameters. This can be visualized in Figure 3, in which both mean functions are plotted.
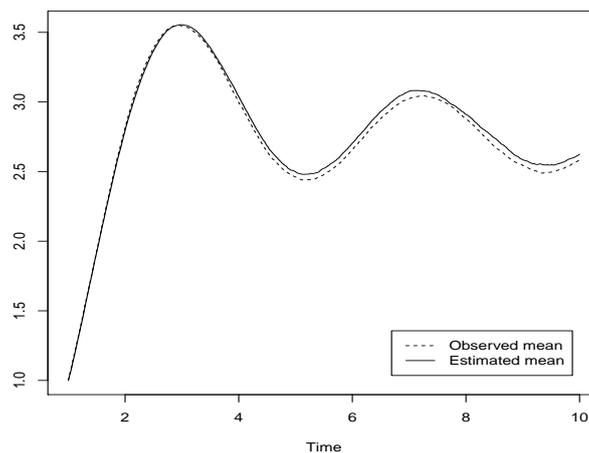

Figure 3. Simulated and estimated mean functions. Gaussian oscillabolastic process.

A new simulation study has been carried out in order to check the process estimation for oscillatory behavior, for example when the parameter β that influences the frequency of the oscillations is varied, and to accommodate for the variability of the sample paths. The simulation pattern is the same as the one considered in the previous case. Parameter β takes values 2, 3, and 4. As previously mentioned, the number of oscillations in the data may determine the range of values of β. Also, the curvature of the trend followed by the oscillations may be a good indicator for the range of parameter θ. Initial values coming from this bounded region lead to the estimation of α and γ.

Final results for different combinations of parameters are shown in Table 2. A graphical comparison between simulated and estimated mean functions, for different parameters, is shown in Figure 4. It is observed that despite the initial difference between the mean function and the original oscillabolastic curve, the Gaussian diffusion process performs well with the proposed methodology for the parameter estimation. Please note that the last two results are the same but for different regions of β_0. Results are the same up to the sign of α and β, due to the odd condition of the sine function.

Also, three simulations has been carried out for different values of the infinitesimal variance, $\sigma^2 = 0.0001, 0.0025$ and 0.01, with parameters $\kappa = 2.3, \alpha = -2, \beta = 2, \gamma = 1$, and $\theta = 0.5$ from $t_0 = 1$ to 10, starting at $x_0 = 1$. Please note that final performance is very good although, as expected, RAE increases with variability. Results are shown in Table 3 and Figure 5.

Table 2. Results for simulated Gaussian diffusion with different parameters.

$\beta = 2$	κ	α	β	γ	θ	σ^2
Original	1.837	-2	2	1	0.500	0.0025
Initial	1.850	-1.792	1.564	0.812	0.561	0.0029
Estimated	1.855	-1.790	1.565	0.813	0.561	0.0029
	$\beta_0 \in [1,3]$		Inflection $t_* = 2.97$		RAE = 0.011	
$\beta = 3$	κ	α	β	γ	θ	σ^2
Original	0.801	-2	3	1	0.500	0.0025
Initial	0.980	-1.942	3	0.611	0.550	0.0013
Estimated	0.977	-1.950	3	0.620	0.550	0.0014
	$\beta_0 \in [2,5]$		Inflection $t_* = 1.98$		RAE = 0.07	
$\beta = 4$	κ	α	β	γ	θ	σ^2
Original	-1.495	-2	4	1	0.500	0.0025
Initial	-1.464	-2.012	4.000	0.836	0.551	0.0020
Estimated	-1.463	-2.012	4	0.838	0.550	0.0021
	$\beta_0 \in [3,5]$		Inflection $t_* = 5.48$		RAE = 0.02	
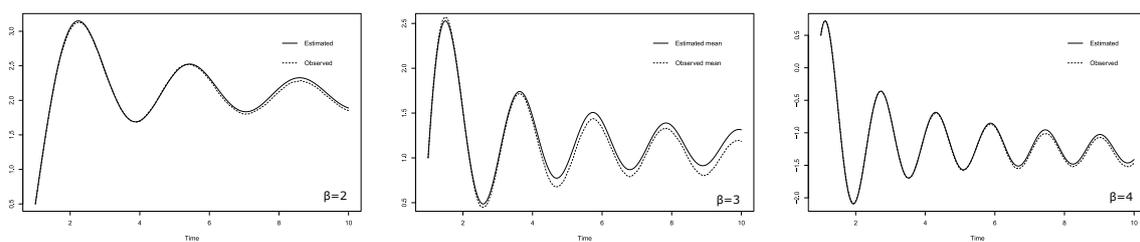
Figure 4. Simulated and estimated mean functions for variations of parameters of Table 2.

Table 3. Results for simulated Gaussian diffusion for different values of σ^2.

$\sigma^2 = 0.0001$						
	κ	α	β	γ	θ	σ^2
Original	2.337	-2	2	1	0.500	0.0001
Initial	2.392	-1.864	2	1.032	0.300	0.0001
Estimated	2.392	-1.869	2	1.038	0.300	0.0001
						RAE = 0.013
$\sigma^2 = 0.0025$						
	κ	α	β	γ	θ	σ^2
Original	2.337	-2	2	1	0.500	0.0025
Initial	2.369	-1.865	2	1.330	0.250	0.0020
Estimated	2.369	-1.868	2	1.332	0.250	0.0022
						RAE = 0.013
$\sigma^2 = 0.01$						
	κ	α	β	γ	θ	σ^2
Original	2.337	-2	2	1	0.500	0.01
Initial	2.479	-1.867	2	0.741	0.300	0.012
Estimated	2.483	-2.059	2.002	0.806	0.304	0.029
						RAE = 0.016
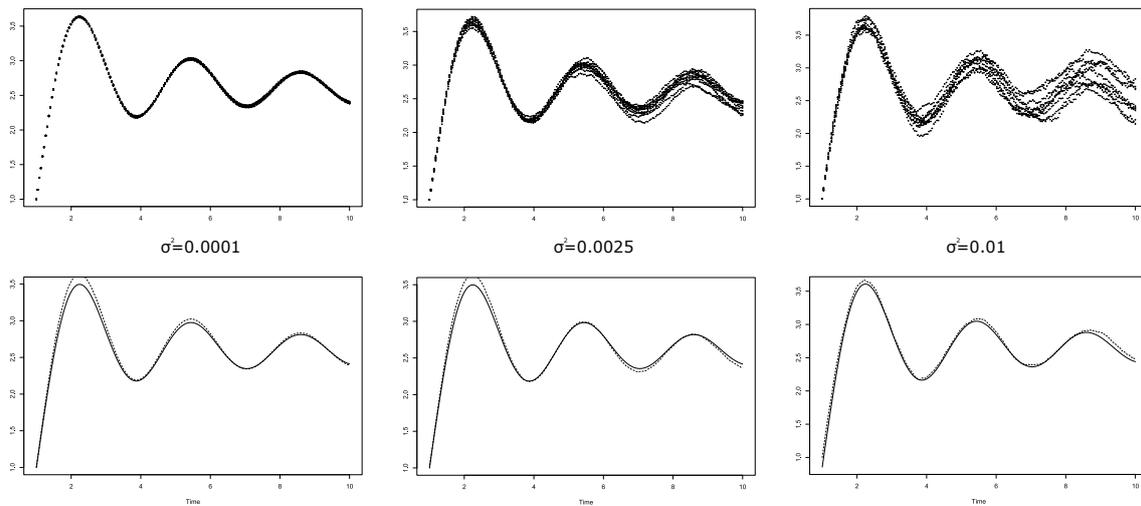
Figure 5. Sample paths (top) and simulated vs. estimated mean functions (bottom) for $\sigma^2 = 0.0001$, 0.0025, 0.01, (see Table 3).

5.2. Simulation in the Lognormal Case

The simulation scheme applied to the lognormal oscillabolastic diffusion process is similar to the one used earlier. In this case, the initial value of the degenerate distribution is $x_0 = 0.5$. Original values for the parameters are $\alpha' = -1, \beta = 2, \gamma' = 1, \theta = 0.5$, and $\sigma^2 = 10^{-3}$. Figure 6 displays the simulated paths (note that in the lognormal case, and in contrast to the Gaussian one, variability grows faster due to its dependence on the state).
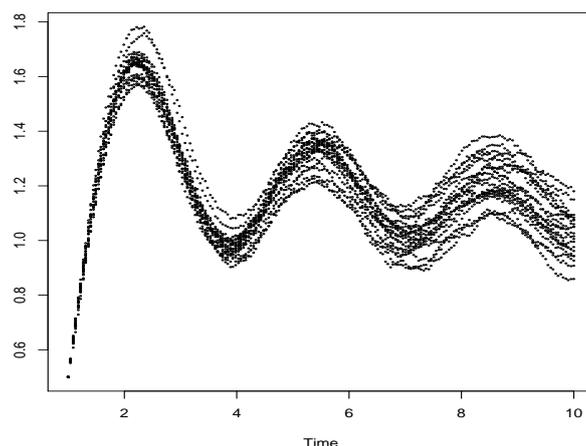
Figure 6. Simulated sample paths for the oscillabolastic lognormal-type diffusion process. Parameters: $\alpha' = -1, \beta = 2, \gamma' = 1, \theta = 0.5$ and $\sigma^2 = 10^{-3}$.

The initial values for the parameters follow from the methodology used in the lognormal case. The search region for β_0 and θ_0 is the same as in the Gaussian simulation example above. The number of oscillations (in relation to the time interval) and the curvature of the observations suggest that this may be a good choice. An inflection condition with $t_* = 2.97$, $u_* = 1.32$ and error $\epsilon = 0.1$ results in initial values $\beta_0 = 2$ and $\theta_0 = 0.55$, respectively. Because of increasing variability, very low error thresholds are not recommended, as they would decrease performance.

Initial values of α' and γ' have been calculated following the same methodology set out in the previous example, resulting in $\alpha'_0 = -0.85$ and $\gamma'_0 = 0.48$. Finally, the initial value of infinitesimal variance σ^2 follows from the regression analysis of the values obtained from the arithmetic and geometric mean of the observed mean values. In this case, $\sigma^2_0 = 0.001$ is obtained.

Using these initial values in the numerical resolution of the maximum likelihood system of Equations (16) and (17) leads to $\alpha' = -0.948, \beta = 2, \gamma' = 0.747, \theta = 0.564$, and $\sigma^2 = 0.001$. Table 4 summarizes the results.

Table 4. Results for simulated lognormal oscillabolastic diffusion.

	α'	β	γ'	θ	σ^2
Estimated values	-0.948	2	0.747	0.564	0.001

Figure 7 shows the simulated and estimated mean functions. Note how the adjustment provided by the mean function is good and is supported by the value of the RAE, which is RAE $= 0.007$.
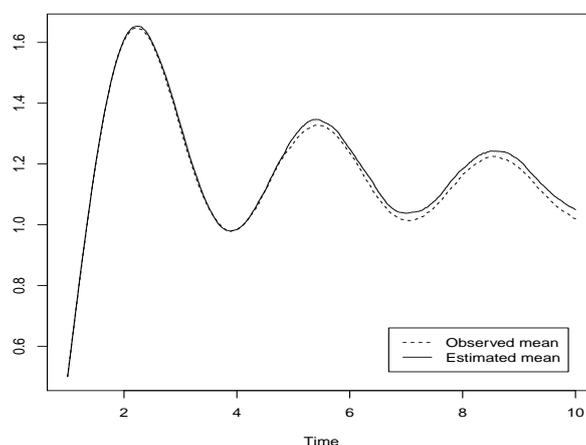

Figure 7. Simulated and estimated mean functions. Oscillabolastic lognormal-type diffusion process.

As for the Gaussian oscillabolastic diffusion process, another simulation study has been carried out by varying, in this case, parameters β (which affects the frequency of the oscillations) and σ^2 (responsible for the variability of the sample paths). Following the same methodology as in the Gaussian case, the ranges of β and θ can be directly proposed from the observations.

In this case, paths are simulated for values of $\beta = 2, 4, 6$. Final estimated values, as well as the search range for initial values of β are shown in Table 5. Furthermore, Figure 8 shows simulated and estimated mean functions for some combinations of parameters. RAE for every one of them is also shown in the table. These values suggest that the model and the proposed estimation methodology perform very well.

On the other hand, Table 6 and Figure 9 also show good performance for different values of variability, in particular $\sigma^2 = 0.0001, 0.0025, 0.01$ for parameters $\alpha' = -0.5.\beta = 2, \gamma' = 1, \theta = 0.5$ and $x_0 = 1$ for $t \in [t_0, 10]$ with $t_0 = 1$.
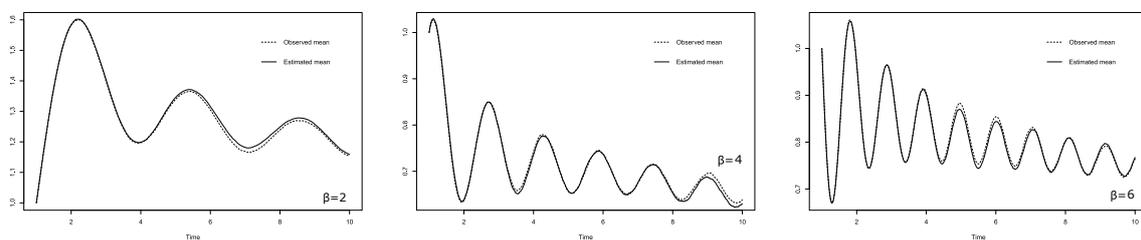

Figure 8. Simulated and estimated mean functions for variations of parameters of Table 5.

Table 5. Results for simulated lognormal diffusion with different parameters.

$\beta = 2$					
	α'	β	γ'	θ	σ^2
Original	-0.500	2	1	0.500	0.0001
Initial	-0.372	2.047	0.353	0.526	0.00002
Estimated	-0.492	1.999	0.813	0.585	0.0001
	$\beta_0 \in [2,5]$	Inflection $t_* = 1.03$		RAE = 0.001	
$\beta = 4$					
	α'	β	γ'	θ	σ^2
Original	-0.500	4	1	0.500	0.0001
Initial	-0.314	4.067	0.196	0.600	0.0002
Estimated	-0.477	3.999	0.694	0.637	0.0002
	$\beta_0 \in [2,5]$	Inflection $t_* = 3.90$		RAE = 0.002	
$\beta = 6$					
	α'	β	γ'	θ	σ^2
Original	-0.500	6	1	0.500	0.0001
Initial	-0.329	6	0.274	0.550	0.0002
Estimated	-0.474	6.000	0.705	0.618	0.0002
	$\beta_0 \in [4,7]$	Inflection $t_* = 0.99$		RAE = 0.004	

Table 6. Results for simulated lognormal diffusion for different values of σ^2.

$\sigma^2 = 0.0001$					
	α'	β	γ'	θ	σ^2
Original	−0.500	2	1	0.500	0.0001
Initial	−0.516	1.789	0.557	0.665	0.00003
Estimated	−0.489	2.001	0.915	0.520	0.0001
					RAE = 0.003
$\sigma^2 = 0.0025$					
	α'	β	γ'	θ	σ^2
Original	−0.500	2	1	0.500	0.002
Initial	−0.498	1.809	0.451	0.667	0.0004
Estimated	−0.498	1.809	0.451	0.667	0.0004
					RAE = 0.004
$\sigma^2 = 0.01$					
	α'	β	γ'	θ	σ^2
Original	−0.500	2	1	0.500	0.010
Initial	−0.506	1.794	0.410	0.674	0.007
Estimated	−0.466	1.979	0.360	0.866	0.006
					RAE = 0.012
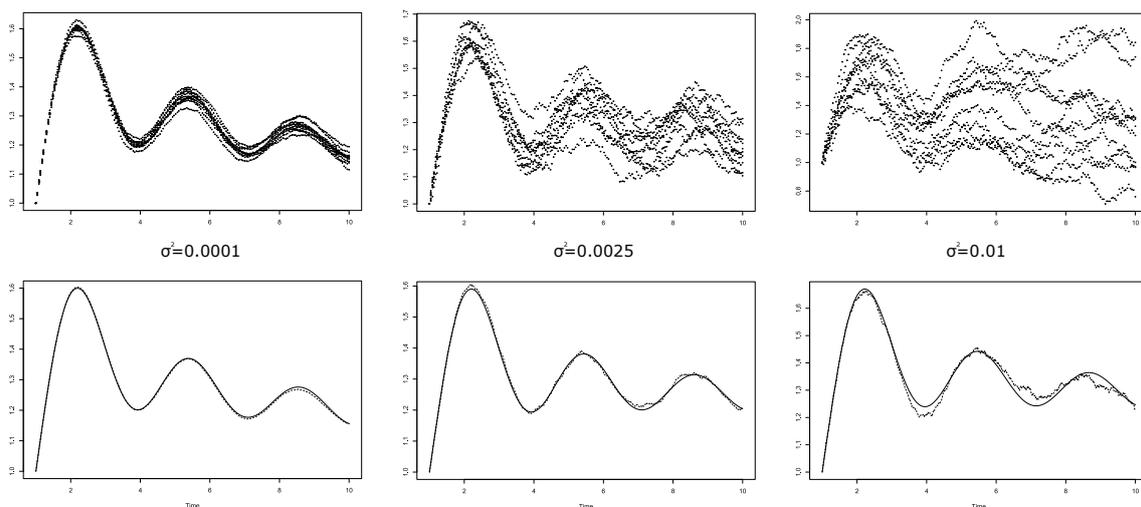
Figure 9. Sample paths (top) and simulated vs. estimated mean functions (bottom) for $\sigma^2 = 0.0001, 0.0025, 0.01$, (see Table 6).

6. Conclusions

Classic growth curves are generally described as arising from ordinary differential equations, which can be studied by themselves or as part of more sophisticated models. In the present paper, we have focused on the study of the oscillabolastic curve, which is particularly interesting for its ability to describe dynamic oscillatory phenomena.

Stochastic models are obtained from deterministic ones by including in them certain random fluctuations that can significantly influence the evolution of dynamic phenomena. This work considers two oscillabolastic stochastic models that retain the properties of the original curve and can be applied to practical situations. Two strategies have been proposed for managing the original differential equation and producing two different stochastic models. One of them is a Gaussian process while the other is a derivative of the lognormal process with exogenous factors.

One of the main problems derived from such models is the development of inference strategies. The estimation of the parameters by maximum likelihood implies the numerical resolution of two systems of equations. We have proposed a methodology to obtain initial solutions for the parameters based on the main traits of the model, namely the presence of an oscillabolastic trend (in the lognormal case) or an asymptotic oscillabolastic behavior (in the Gaussian case). To apply these procedures, a simulation study has been carried out taking into account a variety of error measures. This simulation study confirms the validity of the methodology proposed.

Author Contributions: The three authors have participated equally in the development of this work, either in the theoretical developments or in the applied aspects. The paper was also written and reviewed cooperatively. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Li, W.; Wang, K.; Su, H. Optimal harvesting policy for stochastic logistic population model. *Appl. Math. Comput.* **2011**, *218*, 157–162. [CrossRef]
2. Schurz, H. Modeling, analysis and discretization of stochastic logistic equations. *Int. J. Numer. Anal. Model.* **2007**, *4*, 178–197.
3. Tuckwell, H.C. A study of some diffusion models of population growth. *Theor. Popul. Biol.* **1974**, *5*, 345–357. [CrossRef]
4. Sun, X.; Wang, Y. Stability analysis of a stochastic logistic model with nonlinear diffusion term. *Appl. Math. Model.* **2008**, *32*, 2067–2075. [CrossRef]
5. Hu, G.; Wang, K. The estimation of probability distribution of SDE by only one sample trajectory. *Comput. Math. Appl.* **2011**, *62*, 1798–1806. [CrossRef]
6. Román-Román, P.; Serrano-Pérez, J.J.; Torres-Ruiz, F. Some Notes about inference for the lognormal diffusion process with exogenous factors. *Mathematics* **2018**, *6*, 85. [CrossRef]
7. Román-Román, P.; Serrano-Pérez, J.J.; Torres-Ruiz, F. A Note on Estimation of Multi-Sigmoidal Gompertz Functions with Random Noise. *Mathematics* **2019**, *7*, 541. [CrossRef]
8. Rupšys, P. Understanding the Evolution of Tree Size Diversity within the Multivariate Nonsymmetrical Diffusion Process and Information Measures. *Mathematics* **2019**, *7*, 761. [CrossRef]
9. Rupšys, P. Stationary densities and parameter estimation for delayed stochastic logistic growth laws with application in biomedical studies. *WSEAS Trans. Biol. Biomed.* **2019**, *5*, 117–132.
10. Longtin, A. Stochastic Delay-Differential Equations. In *Complex Time-Delay Systems. Understanding Complex Systems*; Atay, F.M., Ed.; Springer: Berlin, Germany, 2010; pp. 177–195.
11. Sakthivel, R.; Revathi, P.; Ren, Y.; Shen, G. Retarded stochastic differential equations with infinite delay driven by Rosenblatt process. *Stoch. Anal. Appl.* **2018**, *36*, 304–323. [CrossRef]
12. Dung, N.T. Fractional stochastics differential equations with applications to finance. *J. Math. Anal. Appl.* **2013**, *397*, 334–348. [CrossRef]
13. Moghaddam, B.P.; Mendes Lopes, A.; Tenreiro Machado, J.A.; Mostaghim, Z.S. Computational scheme for solving nonlinear fractional stochastic differential equations with delay. *Stoch. Anal. Appl.* **2019**, *37*, 893–908. [CrossRef]
14. Tabatatai, M.; Williams, D.K.; Bursac, Z. Hyperbolastic growth models: theory and application. *Theor. Biol. Med. Model.* **2005**, *2*, 1–13. [CrossRef]
15. Eby, W.M.; Tabatabai, M.A.; Bursac, Z. Hyperbolastic modeling of tumor growth with a combined treatment of iodoacetate and dimethylsulphoxide. *BMC Cancer* **2010**, *10*, 509. [CrossRef] [PubMed]
16. Tabatabai, M.A.; Bursac, Z.; Eby, W.M.; Singh, K.P. Mathematical modeling of stem cell proliferation. *Med. Biol. Eng. Comput.* **2011**, *49*, 253–262. [CrossRef] [PubMed]

17. Tabatabai, M.A.; Eby, W.M.; Bursac, Z. Oscillabolastic model, a new model for oscillatory dynamics, applied to the analysis of Hes1 gene expression and Ehrlich ascites tumor growth. *J. Biomed. Inform.* **2012**, *45*, 401–407. [CrossRef] [PubMed]
18. Tabatabai, M.A.; Eby, W.M.; Singh, K.P.; Bae, S. T model of growth and its application in systems of tumor-immune dynamics. *Math. Biosci. Eng.* **2013**, *10*, 925–938. [CrossRef] [PubMed]
19. Barrera, A.; Román-Román, P.; Torres-Ruiz, F. Hyperbolastic type-III diffusion process: Obtaining from the generalized Weibull diffusion process. *Math. Biosci. Eng.* **2019**, *17*, 814–833. [CrossRef] [PubMed]